DETECTOR

Knowledge-based systems for dairy farm management support and policy analysis

Methods and applications

W.H.G.J. Hennen

STELLINGEN

behorende bij het proefschrift

DETECTOR
Knowledge-based systems for dairy farm management support and policy analysis; methods and applications

door W.H.G.J. Hennen

17 februari 1995

1. De mogelijke verplichting tot het bijhouden van een mineralenboekhouding zal de management-ondersteunende waarde van het fiscaal boekhoudverslag sterk doen toenemen.
   Dit proefschrift.

2. De meeste management informatiesystemen hebben als beoogd doel het verbreden van inzicht en het vergroten van de kennis van de ondernemer. Aangezien deze doelen moeilijk te kwantificeren zijn is het derhalve moeilijk de waarde van zulke systemen te kwantificeren.
   Dit proefschrift.

3. Bij de ontwikkeling en toepassing van kennissystemen voor het analyseren van boekhoudverslagen is de disuniformiteit van verschillende boekhoudsystemen een grotere "bottleneck" dan het proces van kennisacquisitie.
   Dit proefschrift.

4. Voor de melkveehouderij zijn bedrijfsvergelijkende maatstaven zeer geschikt om te gebruiken als referentiewaarde en voor het analyseren van gegevens door experts of kennissystemen.
   Dit proefschrift.

5. Er wordt in veel modellen voor beleidsstudies onvoldoende rekening gehouden met de grote pluriformiteit van bedrijfsituaties, ondernemers en reactiemogelijkheden.
   Dit proefschrift.

6. Wereldwijde uniformering van literatuurlijsten in combinatie met een vermelding van de correcte referentie bij het begin van elk boek of artikel zou voor het werk van onderzoekers eenvoudiger maken.

7. Door de verzelfstandiging van het landbouwkundig onderzoek worden van de onderzoeker naast wetenschappelijke ook commerciële kwaliteiten verwacht.

8. Bij toepassing van de onderzoeksmethode van groepsvergelijking voor het bestuderen van de technisch/economische effecten van een specifieke investering of bedrijfssysteem wordt onvoldoende rekening gehouden met het feit dat ondernemers met een bepaalde grondhouding, niveau van management en capaciteit eerder zouden kunnen kiezen voor de betreffende investering of het systeem.

9. In sommige tijden zijn cashcows waardevoller dan spaarvarkens.

10. Koeien met een té hoge conditiescore hebben een té slechte conditie om goed te presteren.

11. Dat emancipatie bij ministeries nog onvoldoende is doorgevoerd blijkt uit het feit dat het kind van een mannelijke medewerker niet in aanmerking komt voor kinderopvang.

12. De maatschappij is er onvoldoende op ingesteld dat de vrouw na haar huwelijk haar eigen naam wenst te gebruiken.

13. Een flinke heroïnevangst door de politie hoeft niet als een succes te worden aangemerkt indien door de grotere schaarste hieraan de criminaliteit toeneemt.

14. Het beroep van leraar wordt vaak ten onrechte niet alleen financieel ondergewaardeerd.

DETECTOR

KNOWLEDGE-BASED SYSTEMS FOR DAIRY FARM MANAGEMENT SUPPORT AND POLICY ANALYSIS

METHODS AND APPLICATIONS

Promotoren: dr. ir. L.C. Zachariasse,
hoogleraar in de Economische Aspecten van Management en Management Ondersteunende Systemen in de Land- en Tuinbouw

dr. ir. A.W.J. Kolen,
hoogleraar Besliskunde, Rijksuniversiteit Limburg

W.H.G.J. Hennen

DETECTOR
Knowledge-based systems for dairy farm management support and policy analysis; Methods and applications

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
dr. C.M. Karssen,
in het openbaar te verdedigen
op vrijdag 17 februari 1995
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen

ABSTRACT

DETECTOR
KNOWLEDGE-BASED SYSTEMS FOR DAIRY FARM MANAGEMENT SUPPORT AND POLICY ANALYSIS; METHODS AND APPLICATIONS (DETECTOR: KENNISSYSTEMEN VOOR MANAGEMENT ONDERSTEUNING OP MELKVEEBEDRIJVEN EN BELEIDSSTUDIES; METHODEN EN TOEPASSINGEN)
Hennen, Wil H.G.J.
The Hague, Agricultural Economics Research Institute (LEI-DLO), 1995
Onderzoekverslag 125
ISBN 90-5242-273-7
205 pp., tab., fig.

This thesis describes new methods and knowledge-based systems for the analysis of technical and economic accounting data from the year-end records of individual dairy farms to support the management and, after adaptation, for policy analysis.

A new method for farm comparison, the farm-adjusted standard, which makes it able to compare a dairy farm with similar farms, is described. Two methods for the acquisition, representation, and presentation of knowledge from experts are developed. These methods, IMAGINE and FUZZY-DETECTOR, can be used for different types of data and are especially characterised by fuzzy boundaries, compensatory mechanisms, and fast and easy knowledge acquisition.

Two knowledge-based systems have been developed where these methods are used. GLOBAL-DETECTOR performs a global analysis of year-end results concerning aspects of gross margin. ENVIRONMENT-DETECTOR gives suggestions for a reduction of nitrogen surplus while maintaining the income. Proposed is a new method for sector responses on government options (APPROXI). A model is presented as an example of this method.

The developed methods and knowledge-based systems for this thesis offer many opportunities for application in other domains.

Management/Knowledge-based systems/Knowledge acquisition/Accounting/Dairy farming/Economics/Gross margin/Computer software/Fuzzy sets/Farm comparison/Policy analysis/Environmental policy

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Hennen, Wil H.G.J.

Detector: knowledge-based systems for dairy farm
management support and policy analysis : methods and
applications / Wil H.G.J. Hennen. - The Hague :
Agricultural Economics Research Institute (LEI-DLO). -
Fig. tab. - (Onderzoekverslag / Landbouw-Economisch
Instituut (LEI-DLO) ; 125)
With ref.
ISBN 90-5242-273-7
NUGI 835
Subject headings: knowledge-based systems ; dairy farming.

"*For agriculture to make increased use of computer technology, it will find the challenges to be great - but the opportunities will be even greater!*"

(Stephen B. Harsh, 1990)

CONTENTS

Page

WOORD VOORAF 11

SUMMARY 13

SAMENVATTING 19

1. AUTOMATED DECISION SUPPORT FOR DAIRY FARM MANAGEMENT: USEFULNESS AND FEASIBILITY 25
   1.1 Dairy farm management in a changing environment 25
   1.2 The usefulness of supporting dairy farm management 26
      1.2.1 Differences in farms and farm management 26
      1.2.2 Farm management, analysis and critical success factors (CSFs) 28
      1.2.3 Difficulties in analysing farm data 30
   1.3 The possibilities of supporting dairy farm management 31
      1.3.1 Support from advisors 31
      1.3.2 Support from information technology 32
   1.4 Objectives and scope of the study 34
   1.5 Organisation of the thesis 35

2. THE ANALYSIS OF DATA FROM DAIRY FARM ACCOUNTS WITH FARM-ADJUSTED STANDARDS 38
   2.1 Introduction 38
   2.2 Accounting data 40
   2.3 Standards for comparative analysis 41
   2.4 Farm-adjusted standards (FASs) 45
      2.4.1 The method: the construction of the models 46
      2.4.2 The use of FASs 50
      2.4.3 The use of FASs in organisations 53
      2.4.4 Evaluation of FASs 54

3. KNOWLEDGE-BASED SYSTEMS FOR THE INTERPRETATION OF DATA FROM FARM ACCOUNTS 56
   3.1 Knowledge, knowledge transfer, and knowledge-based systems (KBSs) 56
      3.1.1 Knowledge 56
      3.1.2 Knowledge transfer 57
      3.1.3 Knowledge-based systems (KBSs) 58
   3.2 The management of uncertainty in KBSs 61
      3.2.1 Heuristic approach for the management of uncertainty 61
      3.2.2 Fuzzy sets for the management of uncertainty 62

Page

3.2.3 Uncertainty regarding the knowledge of weight and importance 63
3.3 Application of KBSs for analysis of year-end results 64
3.3.1 A short review from the literature 65
3.3.2 Motivation for the development of DETECTOR 66

4. IMAGINE: METHOD FOR ACQUISITION, REPRESENTATION AND PRESENTATION OF KNOWLEDGE IN QUANTITATIVE DOMAINS 69
4.1 Introduction 69
4.2 Model-based knowledge acquisition with IMAGINE 71
4.3 The algorithm of IMAGINE 74
4.3.1 Extended example 79
4.4 The presentation of the conclusions 84
4.5 Explanation facilities: IMAGINE versus rule-based systems 85
4.6 Discussion: prospects and limitations of IMAGINE 86
4.7 Conclusion 87

5. FUZZY-DETECTOR: FUZZY SETS FOR PERFORMANCE EVALUATION UNDER UNCERTAINTY 88
5.1 Introduction 88
5.2 General outline 89
5.2.1 A short introduction to the fuzzy set theory 89
5.2.2 The knowledge base of FUZZY-DETECTOR 91
5.2.3 The use of membership functions (MSFs) in FUZZY-DETECTOR 92
5.3 Matching condition and farm datum 98
5.4 Relevance of the conclusion 104
5.5 The use of FUZZY-DETECTOR in KBSs 106
5.6 Expert's role in FUZZY-DETECTOR 107
5.7 Concluding remarks 108

6. GLOBAL-DETECTOR: KNOWLEDGE-BASED SYSTEM FOR ANALYSIS AND DIAGNOSIS OF PERFORMANCE ON DAIRY FARMS 110
6.1 Some technical specifications of GLOBAL-DETECTOR 111
6.2 The farm data 112
6.3 Analysis of the farm data 112
6.3.1 The farm-adjusted standards (FASs) 112
6.3.2 Graphical presentation of results and effects 115
6.3.3 Analysis of historical data 118
6.3.4 Most striking features 119
6.4 Diagnosis of the farm 120
6.4.1 The method IMAGINE for knowledge acquisition and representation 120

Page

6.4.2 The method FUZZY-DETECTOR as an alternative for IMAGINE 122
6.4.3 IMAGINE versus FUZZY-DETECTOR 125
6.5 Accounting for user's need and management behaviour 127
6.5.1 Investigation of user's need 128
6.5.2 Participation of users in the development 128
6.5.3 Flexibility of GLOBAL-DETECTOR 129
6.5.4 User-friendliness and explanation facilities of GLOBAL-DETECTOR 129
6.5.5 Different styles of farming 130
6.6 Evaluation of GLOBAL-DETECTOR 131
6.6.1 Validation of GLOBAL-DETECTOR 131
6.6.2 Verification of GLOBAL-DETECTOR 132
6.7 How to make use of GLOBAL-DETECTOR 133
6.7.1 Use of GLOBAL-DETECTOR by farmers 133
6.7.2 Interaction with (detailed) systems 134
6.7.3 Transferability to other organisation 135
6.8 Discussion 136

7. ENVIRONMENT-DETECTOR: KNOWLEDGE-BASED SYSTEM FOR REDUCTION OF NITROGEN SURPLUS ON DAIRY FARMS WHILE MAINTAINING INCOME 137
7.1 Introduction 137
7.2 Getting started with ENVIRONMENT-DETECTOR 139
7.2.1 The mineral account 139
7.2.2 The position of the farm 140
7.2.3 The position on other farms 142
7.2.4 Suggestions to decrease the nitrogen surplus while maintaining income 143
7.2.5 Generation of tactics (packages of suggestions) 146
7.2.6 The arithmetical model 147
7.2.6.1 Default values 148
7.2.6.2 Creation of the input list for the second part 149
7.2.6.3 Calculation of effects with the arithmetical model 151
7.2.7 Farmer's own tactic 154
7.3 Concluding remarks 157

8. METHOD TO ESTIMATE SECTOR RESPONSES BASED ON COMPOSING RESULTS FROM INDIVIDUALLY USED KNOWLEDGE-BASED SYSTEMS (APPROXI method) 160
8.1 Introduction 160

	Page
8.2 From ENVIRONMENT-DETECTOR to the APPROXI model	164
8.2.1 The data	166
8.2.2 Alternative options and autonomous developments	166
8.2.3 List of economic measures	168
8.2.4 Styles of farming (attitude)	169
8.2.4.1 Adjustment of the measures	171
8.2.5 Formation of the expected behaviour	173
8.2.6 The arithmetical model: calculation of effects on individual farms	173
8.2.7 Aggregation of results	175
8.3 Results from the APPROXI model	176
8.4 Discussion	179
GENERAL DISCUSSION	183
REFERENCES	191
CURRICULUM VITAE	205

WOORD VOORAF

Gedurende de laatste jaren heb ik met veel plezier en interesse gewerkt aan de ontwikkeling van kennissystemen om melkveehouders te ondersteunen in hun management en aan methoden om kennis van experts te vergaren. Vooral het multi-disciplinaire karakter ervan was boeiend. Tijdens het uitvoeren van het onderzoek en het schrijven van het proefschrift heb ik veel steun en belangstelling mogen ontvangen. Vanaf deze plaats wil ik enkele personen hiervoor graag persoonlijk bedanken.

Die dank gaat in de eerste plaats uit naar mijn promotoren, prof. Zachariasse en prof. Kolen, voor de mij geboden mogelijkheid, voor hun steun, hun inbreng en hun grote belangstelling voor mijn onderwerp. Een woord van dank gaat verder uit naar collega's van LEI-DLO die me hebben geholpen met de ontwikkeling van de methoden en systemen en naar degenen die zeer nuttig commentaar op het proefschrift hebben gegeven. Met name wil ik Tjomme de Haan danken voor de hulp betreffende de door hem ontwikkelde bedrijfsvergelijkende maatstaven en overige relaties en voor de benodigde invoergegevens. Er is dankbaar gebruik gemaakt van de gegevens van praktijkbedrijven in de representatieve LEI-DLO-landbouwsteekproef. Co Daatselaar wordt bedankt voor het beoordelen van de kennissystemen, en Jan-Erik Wien en Willy Baltussen voor de totstandkoming van het APPROXI-model. Een zéér speciaal woord van dank gaat uit naar mijn begeleider Wim de Hoop. Het resultaat is vooral tot stand gekomen dankzij Wim's deskundigheid, motivatie, inspiratie, grote betrokkenheid en opoffering van vrije tijd. Verder wil ik de veehouders A. Berkhout, G. Prins, H. Sikkinga, K. Smidt, J. Terwel en in het bijzonder B. Prins, bedanken voor hun inzet bij de ontwikkeling van de systemen en de uitstraling ervan. Voor de begeleiding van mijn eerste schreden in het interessante gebied der kunstmatige intelligentie ben ik David Brée, Ruud Smit en wijlen Wim Schreinemakers dank verschuldigd. I owe thanks to Stephen Harsh for his interest and his hospitality during my stay in the US.

In mijn dank wil ik ook Corinne Rodenburg, Ton van Lierop en de medewerkers van tekstverwerking (o.l.v. Ria Mannee) en drukkerij van LEI-DLO betrekken. Zij hebben onder tijdsdruk een professionele vormgeving van dit proefschrift gerealiseerd.

Tenslotte wil ik in het bijzonder Julienne danken voor haar inbreng, begrip en steun, en mijn (schoon)familie voor hun belangstelling in mijn onderzoek.

Den Haag, januari 1995 Wil Hennen

SUMMARY

Introduction

Dairy farmers try to achieve their goals. This asks for good farm management, or, the right decisions will have to be made at the right moments. Purposefully gathered and up to date information is indispensable to make this possible. The complexity and importance of farm management has increased because of the increasing complexity of the changing environment, the milk quota system and pollution problems. These ask for better decision-making within the current farm set-up (i.e. tactical management) to achieve low costs.

The quality of management can be improved on many farms because there are great differences in income and in the efficiency of mineral usage between farms under comparable circumstances. A professional analysis of the available (accounting) data is one way to improve management, but this is difficult for farmers because of the lack of required knowledge, good farm comparison methods and good standards for evaluation. So dairy farmers should be supported in analysing, but support from advisors (extension workers) is becoming costly. However, recent developments in information technology give prospects for computers for analysis and diagnosis. According to the literature, especially so-called knowledge-based systems (KBSs) are advocated for this task, despite the low implementation rate of personal computers on farms.

This thesis investigates the possibilities for developing methods and KBSs - according to the requirements (e.g. accounting for the specific situation of the farm, stimulate the creativity of the farmer) - for the analysis of technical and economic accounting data from the year-end records of individual dairy farms to support mainly the evaluation and tactical management functions. Term 'tactical' refers to decisions within the current farm set-up. Since there are differences in farmers' reaction on policy measures, an additional objective of this research is the development of a method (APPROXI) for the evaluation of various government options based on a KBS.

Suitable methods for farm comparison and standards for evaluation, methods for the acquisition of knowledge and expertise regarding data analysis and interpretation, and a guideline for analysis are developed and described in this thesis. The knowledge in the KBSs is predominantly supplied by D.W. De Hoop, and also by C.H.G. Daatselaar and the farmers from our test group.

Data and analysis with farm-adjusted standards

Information is the key element in the decision-making process. Data are the raw material for information. Accounting data to be used by the proposed KBSs should not only be available, but also automatically readable from the data bases of different accountancies. This last aspect is hampered since similar data from different accountancies are at the moment not uniform in description and meaning.

Once the data are available, they can be analysed. In the first step the data must be compared with relevant reference values or standards. In the second step differences between standards and results are judged by the KBS to come to strong and weak aspects and suggestions for improvement.

According to a number of requirements for use in the proposed KBSs of this thesis, five different types of standards are compared: standards based on data from previous years, standards from planning or budgeting, standards as outcome of research (normative), standards from group-comparison, and farm-adjusted standards (FASs). The FASs have many advantages in the research described in this thesis. This new type of standard, which makes it able to compare a dairy farm with other very comparable farms and which is based on regression analysis, is developed by De Haan (1991) and described extensively in chapter 2.

Interpretation of data with KBSs

Knowledge is indispensable for the interpretation of accounting data and for a judgement of deviations from standards. Generally, farmers only incidentally study the account, so it is questionable if they have enough knowledge and experience for this task. Specialists or experts, for example advisors and extension workers, can see things on a farm or on an account in their true perspective. Since only a limited number of farmers makes use of the knowledge from advisors, etc, for the evaluation of their year-end results, distribution by means of KBSs may be of interest.

The knowledge has to be extracted from advisors and experts and transformed to KBSs. The acquisition of knowledge from experts and the representation of that knowledge in a KBS is often difficult and time-consuming. The problems originate mainly from the 'compiled' procedural kind of knowledge that cannot be verbalised. KBSs are defined and some elements and aspects are briefly described. The emphasis is on the management of uncertainty in KBSs, and the difference between the use of a commercial tool and the use of a computer language for the development of KBSs. For this thesis, all tools (i.e. software products based on the methods) and KBSs are developed from scratch with computer language muLISP.

IMAGINE: a method for knowledge acquisition and representation

The quantitative character of the data hampers the development of KBSs in technical, economic, or financial domains. Such data are predominantly continuous and may therefore attain numerous different values. Current tools cannot adequately deal with continuous variables.

A subdivision in several classes, or making these continuous variables discrete, results in an unmanageable number of situations. This problem is called the problem of combinatorial explosion.

To face combinatorial explosion, the Artificial Intelligence method IMAGINE has been developed and described in this thesis. The main characteristics of IMAGINE are fuzzy boundaries, compensatory mechanisms, fast and easy knowledge acquisition and representation, and comprehensive explanation facilities (in contrast with rule-based systems). With this method, cognitive models of the expert regarding conclusions (i.e. strong and weak aspects of management and suggestions for improvement) can be developed. A tool has been built based on this method.

The expert fills in a standard form. He describes which accounting data and FAS values must be used, and how they relate to each other and to the conclusion. Parameters for the different types of parameterised functions are also indicated by the expert. His task is rather easy after some explanation and experience, and he can do this independently from the knowledge engineer (developer of the KBS). The process is called model-based or backward knowledge acquisition.

The contents of the standard form (i.e. the acquired model of the expert), is put in the knowledge base of the KBS (representation). During consultation of the KBS, actual farm data are used as input for the model of the expert. The result (output) is a number indicating the relevance of the conclusion.

IMAGINE is used for the development of some KBSs. The experiences are positive.

FUZZY-DETECTOR: a method for knowledge acquisition and representation

During some tests it appeared that users who are not so well informed about IMAGINE have some trouble understanding the explanation facilities of systems developed with this method. These users urged on the developers the necessity of a less quantitative approach and as a result more clearness. Another limitation of IMAGINE is the disability to deal with uncertain and qualitative data. It is to be expected that such data (e.g. farmer's goals and wishes) become increasingly important in KBSs.

The Artificial Intelligence method FUZZY-DETECTOR has been developed and described in this thesis. This method is an extension of the approach described by Baas and Kwakernaak (1977). Like IMAGINE, fuzzy boundaries, compensatory mechanisms, and the ease of knowledge acquisition and maintenance are the main characteristics of FUZZY-DETECTOR. However, FUZZY-DETECTOR tackles the problem of dealing with qualitative and uncertain data and has a clearer explanation facility for the user. The most important aspect of the method is the management of uncertainty concerning both expert's knowledge and farm data. The fuzzy set theory is lying at the root of FUZZY-DETECTOR.

Knowledge in FUZZY-DETECTOR takes the form of IF-THEN rules. In the IF part, conditions and their importance are stated in linguistic (qualitative) terms; in the THEN part the conclusion is stated. Actual farm data, which may be either qualitative or quantitative, are matched with the conditions of the IF part. The result of these matches are weighed with the importances which leads to the relevance of the conclusion.

GLOBAL-DETECTOR: a KBS for management support

The KBS GLOBAL-DETECTOR has been developed for the global analysis of year-end results (from farm accounts) concerning aspects of gross margin from dairy farms. The requirements that are set up for GLOBAL-DETECTOR originate predominantly from the research by De Hoop et al. (1988). These have been fulfilled to a great extent. The system tries to fill the gaps of the lack of good performance figures and the lack of good farm comparison. With this instrument, the analysis of accounting data by farmers may be improved.

FASs have been developed for returns, variable and fixed costs, and are used to position farm results with respect to results of comparable farms. These standards are also used to show (empirical) relations and the position of other farms by means of a large number of graphical presentations. The deviations with FASs are the clues for good or bad management that are analysed by the artificial intelligence methods IMAGINE and FUZZY-DETECTOR. The user may choose one of these. The result is a list of strong and weak aspects regarding the farm and farm management as well as suggestions for improvement.

A group of six farmers was involved in the development of GLOBAL-DETECTOR. The system appears to be very user-friendly, supports the management of different types of farmers due to it's flexibility, stimulates the creativity, and has extended explanation facilities. The required maintenance is minimal. Validation and verification is discussed.

ENVIRONMENT-DETECTOR: a KBS for environmental management support

The Dutch government aims to reduce and control the nitrogen losses to the environment. One of the instruments of the government that seems to be effective and acceptable is the so called 'mineral account' plus a levy system for unacceptable mineral losses. A mineral account is a statement of flows of minerals resulting in a net surplus of minerals. There are large differences between dairy farms regarding the surplus of nitrogen per hectare. Measures to be taken by the individual farmer to lower the surplus should be farm specific, taking into account the farm's structure and performed management. This led to the development of ENVIRONMENT-DETECTOR, a KBS for the analysis of the nitrogen efficiency (i.e. surplus) on dairy farms and for the generation of global suggestions for a reduction of nitrogen surplus while maintaining the income as good as possible.

FASs are used for the position of the farm compared to similar farms regarding the nitrogen surplus. Farm specific suggestions for improvement are inferred from accounting data, deviations from FASs, the farmer's objective in reducing the surplus, the farm's structure, and the expected outcomes when separate suggestions are applied on the farm. The most important suggestions are combined in two different tactics (packages of suggestions). An arithmetical model is used to calculate the effects of each tactic on the mineral account, on the returns, and on the variable and fixed costs. A farmer who uses the system can modify these tactics and can even develop and evaluate his own preferred tactic. Farmers were involved in the development of ENVIRONMENT-DETECTOR, so information needs and decision behaviour are taken into account to a great extent.

APPROXI: a method for policy evaluation

LEI-DLO makes calculations regarding the economic and environmental effects of policies on farm, regional, and national level. These studies support the government in the decision concerning what policy measure to take. At the moment linear programming (LP), econometric or simulation models are used as the approach for the estimation of reactions of a sector. These approaches are discussed, and a method suggested and proposed by Baltussen et al. (1993a) that tries to combine the strong aspects of LP and econometric models is elaborated in chapter 8. This method is named APPROXI and should be able to account for differences in behaviour of (individual) farmers on various policy options, estimate the effects on the environment and income as a result of that behaviour, deal with new, big or drastical changes, account for technological change and autonomous developments, account for the current situation and farm specific input/output relations, account for strategic aspects (e.g. continuity of farms), use empirical data from (individual) farms (e.g. from representative farms), combine and incorporate knowledge rather easy, provide insight how behaviour and effects are derived, and have a low maintenance.

A model has been developed as an example to illustrate the method or philosophy and is based on most of the requirements mentioned above. The behaviour of a farmer regarding a policy option is estimated with this model for which data from the FADN are used. The behaviour is based on an economic evaluation (using the knowledge base of ENVIRONMENT-DETECTOR) and styles of farming. Results from the research by

Van der Ploeg et al. have been used by J.J.F. Wien for implementation in the model. The arithmetical model of ENVIRONMENT-DETECTOR is used to estimate the effects on the nitrogen surplus and on the income of that individual farm. Individual behaviours and effects are finally aggregated to sector level.

The model requires further validation. The possibility of the development of an APPROXI model according to all requirements has to be proven, especially the accounting for strategic aspects.

Main conclusions

1. The developed methods and KBSs for this thesis have opportunities to support management on dairy farms and to have a better use of accounting data.
2. For the analysis of farm results we must account for the specific situation of the farm. FASs are very suitable for this in the dairy sector, and without FASs the generation of strong and weak aspects regarding the management and suggestions for improvement would have been less easy and straightforward.
3. KBSs must make an objective and economic analysis. They must be flexible and transparant (with explanation facilities) so that farmers with various styles of farming and with different decision and information behaviour can use the same system.
4. The proposed method for sector responses on government policy measures (i.e. APPROXI), which is based on an individually used KBS (i.e. ENVIRONMENT-DETECTOR), is a good alternative for some econometric and linear programming models.
5. Both KBSs may have different kinds of users, and thereby a widespread use in the dairy farm sector with low required support and limited maintenance. They can also be used as a tool for the development of systems in other branches.
6. The development of KBSs need not be time-consuming when suitable methods and tools for the acquisition and representation of knowledge are used.
7. The methods and KBSs have many opportunities for application in other domains.

SAMENVATTING

Inleiding

Melkveehouders proberen hun doelstellingen te bereiken. Dit vraagt om een goed management, dat wil zeggen de juiste beslissingen moeten op de juiste momenten worden genomen. Hiervoor is juiste en tijdige informatie noodzakelijk. De ingewikkeldheid en het belang van het management is toegenomen door de vermeerderde complexiteit van de veranderende omgeving, door het melkquoteringssysteem en door de milieuproblematiek. Om lage kosten te realiseren moeten betere beslissingen binnen de huidige bedrijfsopzet worden genomen (tactisch management).

Op veel bedrijven kan de kwaliteit van het management worden verbeterd omdat er tussen vergelijkbare bedrijven grote verschillen bestaan in inkomen en in de efficiëntie van het mineralenverbruik. Professionele analyse van beschikbare (boekhoud)gegevens is een mogelijkheid om het management te verbeteren, maar dit is moeilijk voor veehouders vanwege de noodzakelijke kennis en de afwezigheid van goede vergelijkingsmethoden en referentiewaarden. Ondersteuning in het analyseren is dus belangrijk. De huidige ontwikkelingen in de informatie-technologie bieden perspectieven voor het gebruik van computers voor analyse en diagnose. In de literatuur worden hiervoor vooral kennissystemen (KBSs) genoemd, ondanks het geringe gebruik van micro-computers op landbouwbedrijven.

Dit proefschrift onderzoekt de mogelijkheden voor de ontwikkeling van methoden en KBSs voor de analyse van de jaarlijkse boekhoudgegevens van individuele melkveebedrijven, teneinde de evaluatie van het bedrijf en het tactisch management te ondersteunen. Dit gebeurt overeenkomstig de gestelde eisen, bijvoorbeeld rekening houden met de bedrijfsspecifieke situatie en stimulering van de creativiteit. Aangezien er verschillen bestaan in het gedrag van boeren op overheidsmaatregelen, is het doel van dit onderzoek tevens de ontwikkeling van een op een KBS gebaseerde methode (APPROXI) voor de evaluatie van verschillende beleidsopties.

Geschikte methoden voor bedrijfsvergelijking, methoden voor het vergaren van kennis betreffende het analyseren en interpreteren van boekhoudgegevens, en een richtsnoer voor analyse zijn ontwikkeld en beschreven in dit proefschrift. De kennis in de KBSs is voornamelijk afkomstig van D.W. de Hoop, alsook van C.H.G. Daatselaar en veehouders van de testgroep.

Gegevens en het analyseren met bedrijfsspecifieke maatstaven (FASs)

Informatie is essentieel in de besluitvorming en gegevens zijn hiertoe de bouwstenen. Boekhoudgegevens die worden gebruikt door de voorgestelde KBSs moeten niet alleen beschikbaar, maar ook direct toegankelijk zijn. Deze toegankelijkheid van databestanden van de verschillende boekhoudkantoren is beperkt door disuniformiteit in de datadefinitie.

Wanneer de gegevens beschikbaar zijn, kunnen ze worden geanalyseerd. In de eerste stap worden de gegevens vergeleken met referentiewaarden. In de tweede stap worden de gevonden verschillen tussen de waarden en werkelijke bedrijfsgegevens door het KBS beoordeeld om te komen tot sterke en zwakke punten van het management en suggesties voor verbetering.

Met betrekking tot de eisen die worden gesteld aan gebruik in de voorgestelde KBSs worden vijf typen van referentiewaarden vergeleken: historische gegevens, planningsgegevens, normatieve gegevens, groepsvergelijking en bedrijfsvergelijkende maatstaven (FASs). Voor het onderhavige onderzoek hebben FASs vele voordelen. Deze op regressie-analyse gebaseerde maatstaven zijn door De Haan (1991) ontwikkeld en in hoofdstuk 2 van dit proefschrift uitgebreid beschreven. Maatstaven maken het mogelijk een bedrijf met gelijksoortige andere bedrijven te vergelijken.

Het interpreteren van gegevens met KBSs

Kennis is essentieel voor het interpreteren van boekhoudgegevens en voor het beoordelen van verschillen met referentiewaarden. In het algemeen bestuderen boeren een verslag slechts incidenteel, en het is de vraag of zij voor het zelf interpreteren voldoende kennis en ervaring hebben. Experts, bijvoorbeeld voorlichters, zien zaken vaak in de juiste verhoudingen. Aangezien slechts een gering aantal boeren voor het interpreteren van boekhoudgegevens gebruik maakt van de kennis van de voorlichting, zou verspreiding van kennis via KBSs interessant kunnen zijn. KBSs kunnen ook een goed hulpmiddel zijn voor de voorlichters.

Kennis moet worden verkregen van adviseurs en experts, en vervolgens worden omgezet in KBSs. Kennisvergaring van experts en het representeren ervan in een KBS is vaak moeilijk en tijdrovend. De problemen worden vooral veroorzaakt door de "gecompileerde" procedurele aard van de kennis welke daarom niet in woorden kan worden uitgedrukt. KBSs worden in dit proefschrift gedefinieerd en enkele onderdelen en aspecten ervan worden kort beschreven. De nadruk ligt op het management van onzekerheid in KBSs, en op het verschil in gebruik tussen een commerciële "shell" en een computertaal bij de ontwikkeling van KBSs. Voor dit proefschrift zijn alle tools (softwareprodukten gebaseerd op de methoden) en KBSs ontwikkeld met computertaal muLISP.

IMAGINE: methode voor het vergaren en representeren van kennis

Het kwantitatieve karakter van gegevens beperkt de ontwikkeling van KBSs in technische, economische en financiële domeinen. Zulke gegevens zijn voornamelijk continu en kunnen derhalve vele verschillende waarden aannemen. De huidige tools kunnen onvoldoende omgaan met deze continue variabelen. Een onderverdeling in verschillende klassen, ofwel het discreet maken van continue variabelen, resulteert in een onbeheersbaar aantal situaties. Dit probleem wordt het probleem van combinatoire explosie genoemd.

Om dit probleem het hoofd te bieden, is de methode IMAGINE ontwikkeld en beschreven in dit proefschrift. De belangrijkste karakteristieken van IMAGINE zijn geleidelijke grenzen, mogelijkheden voor compensatie, snelle en gemakkelijke vergaring en representatie van kennis en beknopte uitlegfaciliteiten (in tegenstelling tot regelgebaseerde systemen). Met deze methode kunnen kennismodellen betreffende conclusies (sterke en zwakke punten van het management, suggesties) van de expert worden ontwikkeld. Een tool is gebouwd op basis van deze methode.

De expert vult een standaardformulier in. Hij beschrijft welke boekhoudgegevens en FASs moeten worden gebruikt, en hoe deze aan elkaar en aan de conclusie zijn gerelateerd. Parameters voor de verschillende functies worden eveneens door de expert aangegeven. Na uitleg en wat ervaring is zijn taak redelijk eenvoudig en kan die onafhankelijk van de knowledge engineer (ontwikkelaar van een KBS) worden uitgevoerd. Het proces wordt model-gebaseerde of "backward knowledge acquisition" genoemd.

De inhoud van het standaardformulier (het verkregen model van de expert) wordt gerepresenteerd in de kennisbank van het KBS. Gedurende de consultatie van het KBS dienen bedrijfsgegevens als input voor het kennismodel van de expert. Het resultaat (output) geeft de relevantie van de conclusie. IMAGINE is gebruikt voor de ontwikkeling van enkele KBSs. De ervaringen zijn positief.

FUZZY-DETECTOR: methode voor het vergaren en representeren van kennis

Gedurende enkele tests is gebleken dat gebruikers die niet zo goed geïnformeerd waren over IMAGINE, moeite hadden de uitlegfaciliteiten te begrijpen in KBSs waar deze methode was toegepast. Zij vonden een minder kwantitatieve benadering, en daarmee meer duidelijkheid, noodzakelijk. Een andere beperking van IMAGINE is de onmogelijkheid om te gaan met onzekere en kwalitatieve gegevens. Het is te verwachten dat zulke gegevens (bijvoorbeeld doelstellingen en wensen van boeren) van toenemend belang worden in KBSs.

FUZZY-DETECTOR, eveneens een methode op het gebied van de kunstmatige intelligentie, is ontwikkeld en beschreven in dit proefschrift.

Deze methode bouwt voort op de benadering van Baas en Kwakernaak (1977). Evenals bij IMAGINE zijn de geleidelijke grenzen, mogelijkheden voor compensatie en het gemak bij het vergaren en onderhouden van kennis, de belangrijkste karakteristieken.

FUZZY-DETECTOR kan echter tevens overweg met kwalitatieve en onzekere gegevens en heeft gemakkelijk te begrijpen uitlegfaciliteiten. Het belangrijkste aspect van de methode is het management van de onzekerheid betreffende expertkennis en gegevens. FUZZY-DETECTOR is gebaseerd op de "fuzzy set" theorie.

Kennis wordt opgeslagen in ALS-DAN-regels. In het ALS-deel worden voorwaarden en hun belang in linguïstische (kwalitatieve) termen opgegeven; in het DAN-deel wordt de conclusie opgegeven. Werkelijke bedrijfsgegevens, kwalitatief of kwantitatief, worden vergeleken met de voorwaarden uit het ALS-deel. De hieruit afgeleide "matches" worden vervolgens gewogen met hun afzonderlijk belang, uiteindelijk resulterend in de relevantie van de conclusie.

GLOBAL-DETECTOR: een KBS voor management ondersteuning

Het KBS GLOBAL-DETECTOR is ontwikkeld voor het op globale wijze analyseren van het saldo van melkveebedrijven op basis van boekhoudgegevens. De eisen die aan het systeem zijn gesteld stammen voornamelijk uit het onderzoek van De Hoop et al. (1988). Aan deze eisen wordt in grote mate voldaan. Het systeem komt vooral tegemoet aan het gebrek aan goede kengetallen en methode van bedrijfsanalyse. Met GLOBAL-DETECTOR kan het analyseren van boekhoudgegevens door veehouders worden verbeterd.

FASs voor opbrengsten en voor variabele en vaste kosten zijn gebruikt om de positie van het te analyseren bedrijf ten opzichte van andere, vergelijkbare, bedrijven aan te geven. Deze maatstaven worden ook gebruikt om de (empirische) relaties en de positie van het bedrijf via een groot aantal grafische presentaties te tonen. De verschillen met FASs, welke aanwijzingen vormen voor goed of slecht management, worden geanalyseerd met IMAGINE of FUZZY-DETECTOR. De gebruiker kan uit deze twee kiezen. Het resultaat is een lijst van de voor het bedrijf relevante sterke en zwakke punten voor het management, welke tevens suggesties voor verbetering bevat.

Een groep van zes veehouders was nauw betrokken bij de ontwikkeling van GLOBAL-DETECTOR. Het systeem blijkt bijzonder gebruiksvriendelijk, ondersteunt het management van verschillende typen boeren door de flexibiliteit, stimuleert de creativiteit en heeft uitgebreide uitlegfaciliteiten. Het benodigde onderhoud is minimaal. Aspecten van validatie en verificatie zijn beschreven in het proefschrift.

MILIEU-DETECTOR: een KBS voor ondersteuning van milieumanagement

De Nederlandse overheid tracht stikstofverliezen te reduceren en te controleren. Een instrument dat effectief en acceptabel kan zijn, is de mineralenbalans met een heffingssysteem voor ontoelaatbare overschotten. De mineralenbalans is een overzicht van de aan- en afvoer van mineralen, resulterend in een overschot. Er zijn grote verschillen tussen melkveebedrijven met betrekking tot het stikstofoverschot per hectare. Maatregelen die de individuele boer kan nemen om het overschot terug te dringen moeten bedrijfsspecifiek zijn en moeten rekening houden met de bedrijfsstructuur en het management. Deze constatering heeft geleid tot de ontwikkeling van MILIEU-DETECTOR, een KBS voor de analyse van de efficiëntie van het stikstofgebruik op melkveebedrijven en voor het geven van globale suggesties voor het verlagen van het stikstofoverschot met zoveel mogelijk behoud van inkomen.

FASs zijn gebruikt voor het bepalen van de positie van het bedrijf betreffende het stikstofoverschot. Bedrijfsspecifieke suggesties worden gegenereerd uit boekhoudgegevens, afwijkingen met FASs, de doelstelling van de boer betreffende verlaging van het overschot, bedrijfsstructuur en verwachte uitkomsten bij toepassing van suggesties. De belangrijkste suggesties worden gecombineerd in pakketten van maatregelen. Een rekenmodel wordt gebruikt om voor elk pakket de effecten op de mineralenbalans en op de opbrengsten en kosten uit te rekenen. Veehouders die het systeem gebruiken kunnen pakketten veranderen, of zelfs hun eigen pakketten samenstellen. Bij de ontwikkeling van MILIEU-DETECTOR zijn veehouders intensief betrokken geweest. Op deze wijze is in grote mate rekening gehouden met de informatiebehoefte en het beslissingsgedrag van boeren.

APPROXI: een methode voor beleidsevaluatie

LEI-DLO maakt in haar beleidsstudies berekeningen voor effecten op economie en milieu van beleidsopties op bedrijf-, sector-, en nationaal niveau. Deze studies ondersteunen de overheid bij het kiezen van maatregelen die genomen moeten worden. Momenteel worden lineaire programmerings- (LP), econometrische- of simulatiemodellen gebruikt als benadering voor het schatten van reacties van een sector. Deze benaderingen zijn bediscussieerd in dit proefschrift en er is een methode voorgesteld door Baltussen et al. (1993a) die de sterke punten van LP en econometrische modellen combineert. Deze methode wordt APPROXI genoemd en hiermee wordt getracht rekening te houden met het verschil in gedrag van individuele boeren op verschillende beleidsopties, wordt het effect op het milieu en het inkomen als gevolg van dat verwachte gedrag geschat, kan rekening gehouden worden met nieuwe, grote en drastische wijzigingen, en tevens met technologische veranderingen en autonome ontwikkelingen. Verder houdt APPROXI rekening met de huidige situatie, met bedrijfsspecifieke input/output relaties en met strategische aspecten (bijvoorbeeld continuïteit). Tenslotte wordt gebruik gemaakt van empirische gegevens van (individuele) bedrijven (bijvoorbeeld steekproefbedrijven), moet kennis tamelijk gemakkelijk gecombineerd en incorporeerd worden, dient inzicht gegeven te worden hoe gedrag en effecten worden afgeleid en het model moet tevens weinig onderhoud vragen.

Een model dat is gebaseerd op de meeste van de hierboven gestelde eisen is ontwikkeld als voorbeeld om de methode of de filosofie te illustreren. Door gebruik te maken van gegevens van het boekhoudnet van LEI-DLO is met behulp van dit model het gedrag van de boer betreffende een beleidsoptie geschat. Het gedrag is gebaseerd op economische evaluatie (middels MILIEU-DETECTOR) en bedrijfsstijlen. Resultaten van het onderzoek van Van der Ploeg et al. zijn door J.J.F. Wien gebruikt voor implementatie in het model. Het rekenmodel van MILIEU-DETECTOR is gebruikt om de effecten op stikstofverliezen en op het inkomen van de individuele boer te schatten. Individuele gedragingen en effecten worden ten slotte geaggregeerd tot sector niveau.

Het model behoeft verdere validatie. De mogelijkheid een APPROXI-model te ontwikkelen welke aan alle eisen voldoet (vooral het rekening houden met strategische aspecten), moet nog bewezen worden.

Belangrijkste conclusies

1. De ontwikkelde methoden en KBSs voor dit proefschrift hebben perspectief voor ondersteuning van het management van veehouders en voor het beter gebruik van boekhoudgegevens.
2. Voor het analyseren van bedrijfsuitkomsten moet rekening worden gehouden met de bedrijfsspecifieke situatie. FASs zijn zeer geschikt voor de melkveehouderij. Zonder gebruik van FASs zou het afleiden van sterke en zwakke punten van het management en suggesties voor verbeteringen minder gemakkelijk zijn.
3. KBSs moeten een objectieve en economische analyse maken. Zij moeten flexibel en doorzichtig (met uitlegfaciliteiten) zijn, zodat boeren met verschillende stijlen en met verschillend beslissings- en informatiegedrag hetzelfde systeem kunnen gebruiken.
4. APPROXI, de voorgestelde methode voor het inschatten van de effecten van beleidsopties die is gebaseerd op een individueel gebruikt KBS (MILIEU-DETECTOR), is een goed alternatief voor sommige econometrische- en LP-modellen.
5. Beide KBSs hebben verschillende soorten gebruikers en kunnen daardoor een wijdverbreid gebruik hebben. Benodigde ondersteuning en onderhoud is gering. De KBSs kunnen tevens voor de ontwikkeling van systemen in andere takken worden gebruikt.
6. De ontwikkeling van KBSs behoeft niet tijdrovend te zijn wanneer geschikte methoden en tools voor het vergaren en representeren van kennis worden gebruikt.
7. De ontwikkelde methoden en KBSs hebben perspectief voor toepassing in andere domeinen.

1. AUTOMATED DECISION SUPPORT FOR DAIRY FARM MANAGEMENT: USEFULNESS AND FEASIBILITY

'The decisions that today's producers must make ... require much more information than ever before' (Barrett et al., 1990)

Dairy farm management nowadays is not only more complex, but also more important than a decade ago, especially due to a changing environment (section 1.1). Since there are differences in management between comparable farms, and consequently also differences in income and mineral usage (section 1.2.1), it is clear that improvement of management is critical. Therefore, it is important to find the strong and weak aspects of management by analysing the farm data (section 1.2.2). Performing such analyses is difficult due to a lack of good performance figures, farm comparison methods, and knowledge (section 1.2.3).

Hence, the way in which data analysis can be supported to improve management should be investigated (section 1.3). From the literature, it can be concluded that so-called knowledge-based systems (KBSs) are candidates for assisting both the farmer and his advisor in this issue (section 1.3.2). This will lead us to the objective and scope of this study (section 1.4), namely investigating the possibilities of developing KBSs to support the analysis of farm data, and hence, to support the management of dairy farms.

1.1 Dairy farm management in a changing environment

Dairy farmers have several goals which they try to realise as good as possible. Some goals are general, like high income, continuity of the farm, enough leisure time, etc, while others are more specific (subgoals). The farmer tries to achieve the confluence of the goals, which might be mutually conflicting and constrained by social, economic and legal factors (Zachariasse, 1990). This asks for good farm management.

Farm management is so comprehensive that it is difficult to define (Boehlje and Eidman, 1984). Webster (1988) defines farm management as 'the process by which resources and situations are manipulated by the farm manager in trying with less than full information, to achieve his goals'. Decisions are goal oriented and the right decisions have to be made at the right moments. Therefore, purposefully gathered and up to date information is indispensable to underpin and improve these decisions by reducing the uncertainty (Alleblas, 1991; Gold et al., 1990). Information is the key to efficient farm management and profitability of the farm business (e.g. Morahan et al., 1989).

Each decision on dairy farms is part of an integrated web of all decisions (Zachariasse, 1990), and is made to achieve the goals. Strategic - or long-term - decisions concern predominantly the set-up and the scale of the farm. Due to the milk quota system, which has been applied for a decade in the EC, enlargement of the scale to improve income has been considerably restricted because of scarcity of milk quota and price. Nowadays dairy farmers have to aim at milking their full quota at the lowest possible costs by keeping an eye upon the total return and costs (De Hoop et al., 1988). Reduction of costs, variable as well as fixed, is of the utmost importance to ensure income (Doluschitz, 1989).

An additional serious pressure on the income in the Netherlands is the environmental problem, which urges a more efficient environmental management on the farms and will also lead to additional costs.

In this new situation, tactical management, or decision-making within the current farm set-up, is crucial to achieve low costs. Also the day-today or operational management (control) has to be performed as good as possible to reach the goals of the farm.

It will be obvious that tactical and operational decisions nowadays are more important and more difficult to make than a decade ago when there were no production and environmental constraints. Complex decision-making increases the probability that wrong choices will be made (Wagner, 1993), and things get even tougher by taking into account the complex government policies, quality demands, technological developments, the increasing flow of data and information, etc. On the other hand, it can be noticed that the level of education and skills of the Dutch dairy farmer has increased during the last years. More and more the dairy farmer will be able to cope with this complexity of decision-making. Improved management skills give him 1) also the opportunity to make better use of advices and management support systems.

1.2 The usefulness of supporting dairy farm management

1.2.1 Differences in farms and farm management

Dairy farming is of great importance for the Dutch economy. In 1993 there were about 30,000 specialised dairy farms in the Netherlands (Van Dijk et al., 1994). These farms differ considerably in income and between farms that are comparable in size there are also differences in gross margin (total returns minus total variable costs) per hectare. This is mainly caused by a diversity in milk quota per hectare (De Haan, 1991). Differences in gross margin still exist between farms comparable in milk

---

1) Only for the author's convenience, the farmer, expert, user, etc will be assumed to be male throughout the whole thesis.

quota per hectare. De Haan (1991) compared for accounting year 1985/86 two groups of dairy farms with on average the same area of land (about 37 hectares), number of cows (about 100) and milk yield (about 6,800 kg fat and protein corrected milk per cow). Both groups therefore had comparable milk quota per hectare (about 18,300 kg). The first group of forty farms (out of 283 specialised Dutch farms) had the highest gross margin of about 340,000 NLG per farm. This group can be considered as the 'best' fifteen percent of specialised farms. The average gross margin of the second group which consisted of 43 farms with a low gross margin (the 'worst' fifteen percent), was more than 80,000 NLG per farm lower than the first group.

Differences in the efficiency of mineral usage can be concluded from a study by Baltussen et al. (1992). They compared a group of farms with a low nitrogen surplus per hectare (the 'best' 25%) with a group that consists of the rest of the farms (the 'worst' 75%). Both groups seemed to have nearly the same area of land and milk quota per hectare. Although the gross margin per hectare also showed no big differences, the gross margin per cow was about 300 to 400 NLG higher on farms where the nitrogen surplus was 150 kg lower ('best' group). Farmers in this latter group required less fertiliser and seemed to have a better feed and grassland management. Baltussen et al. (1993a) concluded from this study that a better mineral management can lead to a higher income of farmers while saving the environment.

The management on dairy farms does not only differ due to differences in structure and size (e.g. Schakenraad et al., 1994), but from what was mentioned above it must be concluded that there are also big differences among farms under comparable circumstances. The differences in gross margin per hectare as reported by De Haan (1991) and the differences in the efficiency of mineral usage as reported by Baltussen et al. (1992), indicated distinct levels of management. Differences in management on dairy farms were also outlined by De Hoop et al. (1988). Daatselaar et al. (1993) gave an overview of the main causes of differences in farm income among dairy farms. Thorough studies on Dutch arable farms by Zachariasse (1974) and on Dutch horticultural holdings by Alleblas (1988), showed that farm results strongly depend on management. The studies also show that nearly each farmer has his own strong and weak management aspects, so nearly each farmer can improve his management.

Differences in farm income and the efficiency of mineral usage stem partly from differences in individual goals that farmers try to achieve (section 1.1) and styles of farming (e.g. Roep et al., 1991). But the quality of the performed management may even be more important. Successful farmers make bad decisions less frequently and do the right things more often (Wagner, 1993). Alleblas (1988) analysed and described the influences of various management aspects on the economic results of horticultural holdings. A list to measure the level of management is presented in his study. Zachariasse (1990) presented a comparable list for judging the quality of farm management.

The question arises how management on a dairy farm can be improved. A farmer can take courses, talk to colleagues, study the literature, etc. On the other hand, the farmer should become conscious of his own management by having insight in his own strong and weak aspects after analysing his own data.

1.2.2 Farm management, analysis and critical success factors (CSFs)

Management is a cyclical process, and consists of the management functions planning, implementation and control (Boehlje and Eidman, 1984). The analysis of results is part of the control or evaluation function of dairy farm management. Information from this function can be used for the planning, and after implementation, the results can be evaluated, and so on. The three functions have to be tuned carefully. Figure 1.1 shows a slight part of the general structure of an information model for dairy farming. Such an information model 'describes the functions, the processes, the information flows and the data, which are all important for the management of the farm' (De Hoop, 1988).

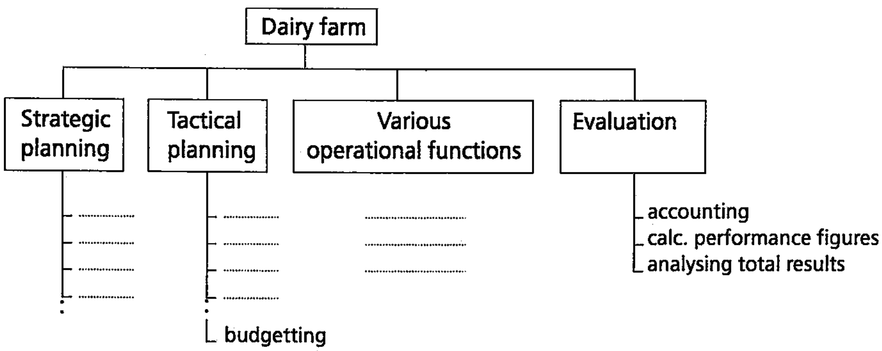


Figure 1.1 Functions and processes of a dairy farm (source De Hoop, 1988)

This figure shows that the management function *evaluation* is split up into three processes: accounting, calculating performance figures and analysing total results. Evaluation of the farm is denoted by dairy farmers as an important function of management (De Hoop et al., 1988). For this function reliable and extensive data are required from accounting and other sources. Such data, and especially data from accounts, are predominantly historical and not oriented to the future. Performance figures have to be calculated from these data and standards have to be set up to find the strong and weak aspects of the realised (last year's) management by analysis.

The analysis can be improved by focussing on the main factors determining the outcome of farming. Such factors can be named Critical Success Factors (CSFs), or 'the few key areas where 'things must go right' for the business to flourish and for the manager's goals to be obtained' (Bullen and Rockart, 1981). Some CSFs are quite general and are relevant for most farms, while others are very farm specific, and depend on the structure of the farm, its situation and the farmer himself. In fact, each farm has its own set of CSFs (Schakenraad et al., 1994). An analysis of CSFs can also identify the major information needs (King et al., 1990).

The analyser can be an advisor, a farmer, etc. Since the task of analysis may be quite difficult for several farmers (see section 1.2.3), we presume for the moment that the analyser (or 'expert') is an advisor or accountant. Such a person could have gained knowledge and experience from farm comparison and from the identification of CSFs in specific situations. When the farmer's goals are known to him, he can analyse farm data and identify the CSFs based on the variables and the farm's values. Else, an economic goal (high income) is assumed and the analyser's task can be performed at his desk in the office. The analyser makes use of his *cognitive model*, where farm data and the calculated performance figures are matched against the CSFs, to come up with the relevant strong and weak aspects of the farm and the management (figure 1.2). A farmer should then try to at least maintain strong aspects, because it is his strength, and weak aspects are eligible for improvement. The cognitive model of a human analyser can be made artificial by a KBS, as we shall see later.

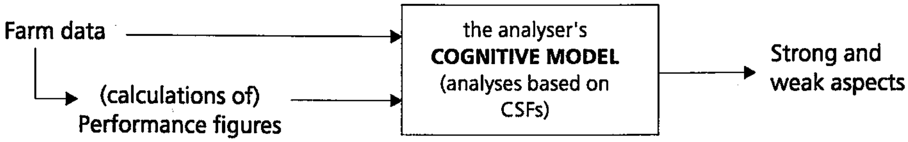


Figure 1.2 Knowledge-based analysis of farm data

The outcome of this analysis delivers useful information for the planning function of the management cycle (figure 1.1). When CSFs and (farmer's) expectations are known, tactics for the next period of time can be formulated. This is followed by the identification of the relevant variables and their target values. After implementation, planned and realised results can be compared and analysed.

1.2.3 Difficulties in analysing farm data

Poppe (1991) comprehensively discussed the use of farm accounts and concluded that accounting and the use of data from accounts is not so popular among farmers. He blamed it on the characteristics of agriculture, such as uncontrollable production processes in small holdings on the one hand, and the contents of the records on the other hand. Since the characteristics of agriculture cannot be changed, Poppe (1991) suggested to alter the contents to improve the information value of farm accounting. For example, inclusion of more technical and decision-oriented data, integration of planning, provision of interim results and alternative presentations of data (layout, graphics, etc). At the moment, records do not produce adequate data for complete business management, especially the financial records for tax reporting purposes are inadequate (Palmer, 1992; Poppe, 1991).

Poppe (1991) could not find any literature to verify the conclusion that the use of accounts would lead to better results, an assumption frequently made. Interesting is, however, that although 'the keeping of accounts as such did not seem to be correlated with farm results ..., the evaluation afterwards is a crucial element. Without this evaluation the influence on farm return is zero' (Poppe, 1991, referring to the research done by Alleblas, 1988). With the word *evaluation* in the quotation is meant the analysis of accounting records.

Farmers find it difficult to analyse farm data from accounts (Dobbins, 1989), and most farmers are therefore unable to fully utilise farm records for farm management. From a study on nearly two dozen farms, De Hoop et al. (1988) concluded that there is generally a lack of good performance figures and methods to compare a farmer's own results with comparable farms. There is little information from other farms which would give farmers a basis for comparison. The need of good standards for farm comparison was also stressed by Phillips and Harsh (1987) and Dobbins (1989). Chapter 2 elaborates on different standards.

From figure 1.1 it shows that calculation of performance figures (and standards) has to be fulfilled before (good) analysis of the results can take place. Since analysis is so difficult, the process requires knowledge from an 'expert' (e.g. advisor, accountant) for several reasons (Dobbins, 1989):

1. what is acceptable for one measure may well depend on the value of another;
2. the interpretation of performance measures may also be affected by events that were beyond the control of the farmer;
3. performance measures are not equal.

In many cases, farmers do not have adequate expertise for such an unaccustomed task. Farmers do not analyse accounts very frequently, year-end results perhaps only once a year, resulting in little experience

(Phillips and Harsh, 1987). The criticism that farmers are unwilling to spend time on analysis or acquisition of knowledge to do so (Smith, 1989; Wagner, 1993), should *not* suggest that the farmer is to blame. Farmer's unwillingness may mainly be caused by the factors discussed earlier in this section, like the lack of good performance figures and methods.

We must draw the conclusion that (dairy) farmers require support in analysing their farm data to detect their strong and weak aspects. On the one hand, analysis is extremely important in a complex environment with so much emphasis on improvement of the management to reduce costs (section 1.1). On the other hand, farmers have difficulties in performing a professional analysis on their own (this section), especially because of a lack of good performance figures and farm comparison methods.

1.3 The possibilities of supporting dairy farm management

Requirements to support dairy farm management, and especially to support analysis of data from accounts, are:

1. adequate technical and economic data from the dairy farm.
2. technical and economic data from other dairy farms for comparative analysis.
3. good methods for farm comparison and standards for evaluation.
4. knowledge and expertise of data analysis and interpretation.
5. a guideline for analysis, where data, methods and knowledge are brought together.

A handful of accountancies, LEI-DLO included, keep technical and economic records from many dairy farms. They are able to supply the data for farm comparison. Unfortunately, the different organisations have different coding systems, different data definitions, etc, which makes the use of comparative data from other organisations rather difficult (chapter 2).

The research at issue will only go into the methods and standards, the knowledge, and a guideline for analysis (points 3-5). Data (points 1 and 2) and the organisations who supply them shall be considered as given and no suggestions will be made for alterations, improvements etc. The starting point is the current flow of technical and economic data from organisations that dairy farmers have access to.

1.3.1 Support from advisors

At the moment the knowledge and expertise for analysing farm results stems from advisors (e.g. extension workers). There is a growing interest for accountancies and other private companies to give their own advice, since an individual consult from the public extension service is no longer free of charge in the Netherlands due to decreasing public funds for extension and because this extension service has not a free or easy access to the data. However, accountancies need the knowledge and the people to give advice. This has to be kept as self-sufficient as possible. Not only financial advice, but especially economic/technical advice is required for management support. Such organisations may need additional, specialised technical expertise that the improvement of farm management asks for, especially when dairy farming is becoming increasingly complex. Advisors might also require supportive knowledge from other areas and good standards for evaluation. So there are organisational bottle necks in supporting farmers.

A solution to increase the support from advisors can be that the extension service and the accountancies work together, or that there will be an easier access to the data, or that the accountancies expand their service with economic/technical analysis and advices.

1.3.2 Support from information technology

Recent advances in computer hardware, software and telecommunications technology have increased the potential for effective computer-based support of farm management decisions (King et al., 1990). Dobbins and King (1988) suggested that a computer can be used as a means to provide a wider access to the expertise of the specialists for the analysis of year-end results. They refer to computer programmes which are called expert systems or KBSs, computer programmes which contain knowledge or expertise and do the job the same way as humans do. Chapter 3 is devoted to the explanation of such systems. KBSs for analysis are also advocated by several other authors (e.g. Phillips and Harsh, 1987; Folkerts, 1989; McGrann et al., 1989; Wagner, 1993). Smith (1989) noted significant opportunities of KBSs for the evaluation of farm management data by using the knowledge and skill available from many dairy extension specialists together in one system. Since dairy management has been increased in complexity, Doluschitz et al. (1988) saw in the use of KBSs promising advantages especially by fusing knowledge and the transfer of research information to levels applicable to end users.

Computers may be favourable in the sense that they can do the task in a combined effort, such as calculating standards, using data from comparable farms, performing comparable analysis, using knowledge from experts, and performing the whole process of analysis according to a certain guideline. Such combined systems, or the combination of KBSs and traditional systems, are often called hybrid systems. Wagner (1993) found hybrid systems particularly useful in situations requiring economic judgements.

It must be emphasised that, according to De Hoop et al. (1988), the role of the human decision maker and advisor cannot be fully taken over by a computer programme. A knowledge-based or hybrid system must merely be seen as an aid in decision-making and as an intelligent assistant for the farm advisor. In dairy farming, the implementation rate of personal computers is 'surprisingly' low (Zachariasse, 1991). Especially the availability of good management information systems is neccessary to make the purchase of a personal computer worthwhile. Although many causes of low adoption are identified in several researches (De Hoop et al., 1988; King et al., 1990; NRLO, 1991; Klink, 1991; Leeuwis, 1993), I will predominantly restrict myself to and link up with those described by De Hoop et al. (1988) because of the relevance for my study. They found that systems:

- do not sufficiently account for the farm's specific situation and the information need and decision behaviour of the manager;
- lack good performance figures and methods;
- have a prescriptive character, but farmers favour a more descriptive system;
- do not give enough insight in the working of the system and the way calculations are performed;
- have no easy access to the data.

King et al. (1990) stated that inadequate understanding of managerial processes and information needs of the farmer might be the greatest impediment to adoption. This understanding can be met by the determination of CSFs. This method has been applied by De Hoop et al. (1988) on nearly two dozens of dairy farms.

Because the implementation rate of personal computers on farms is so low, the effort to let many dairy farmers have access to the knowledge and to the results of a professional analysis should not only come from computer programmes installed on the farms. The proposed KBSs in chapters 6 and 7 are also not limited by the low implementation rate. The alternative is that the programmes run on accountancies or on extension offices and make use of accounting data from a central data base. Accountants can then use them to make an individual analysis report and send it with the account to the farmer. If the farmer needs additional information, then a more detailed analysis can be performed during a visit with his accountant or advisor (e.g. by using advisor's portable computer on the farm).

KBSs can have advantages for an organisation (accountancy or extension). They may increase overall performance, competitive advantage, quality of the product (an account), standardisation, and the professional image of the organisation (Quartel et al., 1992).

Computerised support of the whole tactical, and in particular the operational management, might lead to an extensive system which is hard to manage. A guideline should at first instance efficiently point to the global problems by analysing the farm as a whole. From the detected strong and weak aspects, a detailed analysis on operational level can be performed successively. Such a top-down approach makes efficient search strategies for potential problems possible.

1.4 Objectives and scope of the study

We can conclude that there is a need to support dairy farm management by means of KBSs embedded in hybrid systems. The research at issue will only relate to the analysis of year-end results for mainly the tactical management support, or the first part of the top-down analysis. In this study, tactical management concerns decision-making within the current farm set-up. Here it is assumed that this management can be evaluated by the analysis of year-end results.

Neither the analysis of the financial position nor of the financing of the farm shall be included in this study. Such a kind of analysis will be called *financial analysis* from now on. Although most applications of KBSs for analysis are oriented towards financial analysis (Webster, 1988), the research at issue will cover all aspects of gross margin and to a much lesser extent the fixed costs. This is done because cost reduction is important, and because the internal management of dairy farms is primarily focussed on efficient production, expressed in the gross margin (Zachariasse, 1990; De Hoop et al., 1988).

The required technical and economic data from year-end records are present in the data bases of accountancies. What has to be developed are:

- good methods for farm comparison and standards for evaluation;
- methods for the acquisition of knowledge and expertise regarding data analysis and interpretation;
- KBSs which contain a good guideline for analysis, where data, methods and knowledge are brought together. The KBSs should be able to give suggestions for improvement (advices).

This has led to the following objective of this thesis:
Investigating the possibilities for developing methods and KBSs - according to the requirements stated below - for the analysis of technical and economic data from individual dairy farms to support mainly the evaluation and tactical management functions. As specific application, the use of a KBS for policy analysis.

Requirements 1) for the development of methods and KBSs are:

- use data from year-end accounts which are already available at accountancies (stored in the data bases);
- aim at tactical decision-making (analysing the aspects of gross margin) and the giving of suggestions for improvement;
- take into account the farmer's specific situation (e.g. set-up of farm), information need, decision behaviour, wishes, etc, as good as possible. Important are the possibilities to compare farm results with results of comparable other farms, and the user's involvement during the whole development phase;
- give much insight (descriptive instead of prescriptive), resulting in learn-effects;
- stimulate creativity of the farmer;
- perform easy and fast maintenance (also concerning methods and knowledge);
- deal with qualitative, unreliable, uncertain and missing data;
- advocate widespread use (user friendly, minimum support required, low price, no additional hardware and software);
- apply methods and tools in other domains (e.g. systems for analysis in horticulture or regarding systems outside agriculture) and for other purposes. The methods must also be applicable at a more detailed level.

1.5 Organisation of the thesis

For a good and easy implementation of farm comparison, there is need for an automatic access to data, uniformity of data, good performance figures and good methods for comparison. The accounting data that are necessary for the proposed KBSs, are discussed in chapter 2. Much attention is also given to the aspects of the uniformity of data, since different organisations have different definitions. Several methods or types of standards for analysis are briefly described and compared. The method of farm comparison, called the farm-adjusted standard (FAS), is of special interest. This method was first described by De Haan in 1991 and will be presented and discussed here.

Chapter 3 explains some aspects of knowledge, how to transfer the knowledge from an expert to a KBS, and how KBSs are generally structured. This chapter deals with the aspect of uncertainty at great length. Motivation for the use of KBSs for the analysis of year-end results is given, especially based on a review from literature.

When many performance figures have to be analysed, we face the problem of combinatorial explosion when many continuous variables are

---

1) A number of these requirements are expressed by dairy farmers themselves, according to the study from De Hoop et al. (1988).

made discrete for the sake of analysis. A new method and tool, IMAGINE, has been developed to cope with this problem. This method is presented in chapter 4. IMAGINE deals successfully with the problem of combinatorial explosion by the introduction of smooth or fuzzy boundaries and the possibility of compensation between different data or information sources. With IMAGINE, a cognitive model can be developed that transforms farm data to strong and weak aspects (figure 1.2).

FUZZY-DETECTOR, a new method and tool which tackles the problem of dealing with qualitative and uncertain data, is described in chapter 5. It is to be expected that such data types become increasingly important in (knowledge-based) computer programmes of the future. The method is based on the fuzzy set theory 1), which is (shortly) described in the chapter as well. The way FUZZY-DETECTOR deals with the management of uncertainty is of special interest. The rather theoretic descriptions of FUZZY-DETECTOR in chapter 5, and of IMAGINE in chapter 4, are illustrated with simple examples.

GLOBAL-DETECTOR, the KBS for the analysis of year-end results concerning aspects of gross margin from dairy farms, is presented in chapter 6. Farm-adjusted standards (FASs), and the methods IMAGINE and FUZZY-DETECTOR are applied in this system, as well as a guidance for the analysis. The same methods are used in the system ENVIRONMENT-DETECTOR 2) (chapter 7), a KBS which extracts suggestions on how the dairy farmer can reduce the nitrogen pollution on his farm, while maintaining the economic performance (gross margin) as good as possible. Whereas GLOBAL-DETECTOR is mainly a system for supporting the analysis process of the evaluation function (figure 1.1), ENVIRONMENT-DETECTOR supports also the tactical planning 3) (i.e. budgetting, see also figure 1.1), since the farmer is able to simulate different alternatives for expected effects.

One of the requirements for developing methods and tools is the usage for different purposes, thus not only for usage in the management support systems as described in chapters 6 and 7. In chapter 8, a method for policy analysis on a sectoral level is described, which uses the system ENVIRONMENT-DETECTOR and the tool FUZZY-DETECTOR. This method is called APPROXI 4) and approximates the reactions of a sector on alternative policy or economic options based upon the estimated reactions of individual farms (Baltussen et al., 1993a).

---

1) The term fuzzy set will be explained in chapters 3 and 5.
2) The Dutch name of this system is MILIEU-DETECTOR.
3) Although most suggestions are made within the current farm set-up (e.g. lowering amount of fertiliser or amount of concentrates), there are also suggestions on a more strategic level (e.g. purchase of milk quota).
4) The initial idea behind the method **APPROXI** (**APP**roximation of **R**eactions of various **O**ptions based upon farms **X_i**) was presented by Baltussen et al. (1993a).

Concluding remarks are given in chapter 9, with special emphasis to the prospects and limitations of the proposed methods, tools and systems regarding their usage for individual farm analysis as well as analysis of the sector. The final chapter will end with a proposed research agenda with respect to the study at issue.

2. THE ANALYSIS OF DATA FROM DAIRY FARM ACCOUNTS WITH FARM-ADJUSTED STANDARDS

'Vergelijken onder gelijke omstandigheden vind ik het belangrijkste ...' 1)
(Interview with dairy farmer A. Berkhout in 'Boerderij', VE77-2, 1992)

'Der zwischenbetriebliche Vergleich ist ein sehr wertvolles Instrument zur ermittlung der Quellen des Erfolgs oder Miserfolgs eines Betriebszweigs. Er stellt für die meisten Betriebsleiter einen starken psychologischen Anreiz dar: Es ist nur allzu menschlich, daß jeder Landwirt seine Position im Vergleich zu den anderen erkennen möchte.' 2) (Heinrich und Walter, 1989)

2.1 Introduction

To support dairy farm management, the farmer must actually get help in the decision-making process 3). Information is the key element in this process (e.g. Morahan et al., 1989), and it has value since it decreases uncertainty and it changes the probabilities attached to the expected outcomes in a decision situation (Davis and Olson, 1985). Information is also valuable to gain insight and to increase farmer's knowledge which might be useful for future decisions. Especially this utility makes the value of many information sources and systems in fact non-quantifiable.

Data is the raw material for information: 'Information is data that has been processed into a form that is meaningful to the recipient and is of real or percieved value in current or prospective actions or decisions.' (Davis and Olson, 1985).

Accounting data are readily available as a result of the *accounting process* from the *evaluation function* (see figure 1.1 in chapter 1). This

---

1) "Comparison under equal circumstances is most important to me... (transl.WH)".
2) "Comparison among farmers is a very valuable instrument to find the causes of good or bad performance of a branch. For most managers this comparison conveys a strong psychological attraction: it is just nothing but human that each farmer wants to know his position compared to others (transl.WH)".
3) This process will not be described here. See e.g. Davis and Olson (1985) on the decision making process.

task is mostly carried out by accountancies. For the application of knowledge-based systems (KBSs) concerning the analysis of data from year-end results, these data must not only be available, but also automatically readable from the data bases of accountancies. Furthermore, it would be very favourable if similar data from different accountancies are uniform in description and meaning. The absence of this aspect must be regarded as a serious problem at the moment. Section 2.2 deals with the required accounting data and possible solutions to the uniformity problem.

Once the data are available, they can be analysed. In the first step the data must be compared with relevant reference values or standards. The following *requirements* for a good, effective and efficient analysis are necessary for a method that comes up with the reference values or standards:

1. if relevant, the effort for the development of the standards and the successive maintenance of their algorithms must be as low as possible. The calculation must be an easy task for the accountancy. The interval between finishing the account and such a calculation must be minimal;
2. there must be a good basis for the successive deduction of strong and weak aspects. Differences with standards should reflect or clarify the management of the farm. Standards should therefore account for year effects (e.g. weather, prices) as good as possible;
3. standards must account for the farmer's situation, information need and decision behaviour. They must be farm-specific, of high interest for the farmer and they should motivate him. Standards that might be used as targets should be realistic and attainable, without frustrating the farmer;
4. the standard must be understandable and knowledge about the way it is calculated must be easily transferable.

Section 2.3 describes different methods for farm comparison. In fact, one can also speak about different methods for the calculation of standards. These methods are judged in section 2.3 by the requirements mentioned above. They have considerable limitations for detecting strong and weak aspects. This is a problem for a sound analysis. The method of farm-adjusted standards (FASs), which will match the requirements for the most part, is suggested as an alternative and is described at length in section 2.4. This method is used in the KBSs of chapters 6 and 7.

The deviation between the standards and the realised data from the accounts, are information sources that are useful for current and prospective management decisions and for gaining insight and knowledge. The analysis of data with the aid of these standards will not be discussed in this chapter since they are of concern in the remaining chapters of this thesis.

2.2 Accounting data

In the Netherlands, farmers predominantly make use of the services of accountancies to do the accounting, since farmers are often more interested in the biotechnical aspects of their business (Poppe, 1991). Most Dutch dairy farmers keep only records of those data that are required for tax purposes. These records have limited potential for the production of adequate information for complete business management (Palmer, 1992). They miss essential technical and economic data and the accounts are based on fiscal concepts. King et al. (1986 as well as 1990) stated that the high costs of collecting, organising and entering data, forces most farmers to collect only data for the financial statements and tax returns.

But farmers who make or receive additional farm-economic accounts that are based on economic concepts, have more analytical possibilities. Data from these accounts can roughly be divided into three categories: financial data (e.g. solvency), economic data (e.g. gross margin per cow) and technical data (e.g. milk yield). The environmental data, which are in fact a kind of technical data (e.g. nitrogen surplus per hectare) are of growing interest. In the near future, each dairy farmer will have to administrate the data required for the Manure accounts and/or the Mineral accounts (Poppe, 1992; CLM et al., 1991), leading to extra work and costs for the farmers (Baltussen et al., 1993a). With these additional environmental data, the obligatory financial records for tax purposes can be enriched in such a way that analysis has more management supporting value. Such extended accounts and farm-economic accounts will grow towards one another, eventually resulting in more possibilities and in a large-scale utilisation of methods and computer programmes for analysis.

Accounting data required for the development of the systems as described in this thesis are a handful of economic, technical and environmental data as are available from the accounts of the accounting department of LEI-DLO (Farm Accountancy Data Network, FADN). The limited number of required data limits the depth of analysis, as was intended. The systems are therefore global and integral; detailed analyses are not performed. The concepts, methods and systems in the following chapters are *not* restricted to application in the accounting department of LEI-DLO, but are or should be applicable in other accountancies as well.

We must strive at using accounting data directly from the data base of an organisation without any manual data entry. A current drawback however for a large scale potential of computer programmes is the diversity in data definitions and calculation concepts applied by the different accountancies. In the Netherlands, the method 'Information Engineering' is used in all branches of agriculture to reach uniformity in data definition (see e.g. Zachariasse, 1990; De Hoop, 1988; Poppe, 1991). These data definitions are described in the Data Model of the so-called Information Model. The other part of the Information model is the Process Model, which describes the decision processes of a farm and uses data from the Data Model. LEI-DLO has taken part in the identification of common data requirements and decision processes across farms to make such an Information Model for all financial decisions that are made by farmers. Among other things, this led to the detailed Information Model of the cluster 'Analysis and Diagnosis': a detailed description of the processes and data that are required for gaining an impression of strong and weak aspects of management (LEI/VLB, 1989).

According to Poppe (1991), the uniformity of the terminology is one of the main attractions of using Information Models. But since the description is not enough with respect to accounting, LEI-DLO and the Organisation of Agricultural Accountancies (VLB) have made uniform directives for the use of accounting data by means of a loose-leaf edition with a uniform scheme of account names (chart of accounts). This is named GRAS (Poppe, 1991).

Despite of all these efforts, the use of uniform data and calculations by different organisations has not been fulfilled yet, although some favourable developments have been established. Due to this problem of uniformity, computer programmes (e.g. for the analysis of accounting data) which are especially developed for one accountancy, may not be directly usable by another accountancy. At best a data-transformation programme can be built to harmonise the data and to make several necessary adjustments in the analysis system, so that the uniformity problem can be overcome. With this procedure we take away the symptoms but not the real causes.

2.3 Standards for comparative analysis

Farm results can be compared with standards or reference values. The difference between a standard and a result is an important datum for the judgement in the analytical process, to come to strong and weak aspects. These differences show if factors are high/low or favourable/unfavourable, and in general they are indications of performed management. The kind of standard must be clear to the analyst to make the judgement sound.

References are indispensable in decision-making, and each decision requires a specific set of reference values. Such values can remove part of the uncertainty that surrounds decision-making. There are actually many different standards or reference values, some are quantitative and can be calculated from accounting data, while others are very qualitative and rather vague. A conversation with colleagues can adjust the reference pattern in the mind of the farmer and he can extend his knowledge by learning. The attitude and interest for reference values is very personal; one farmer likes to compare his results with the results from last year, while the other is more interested in the results from colleagues. This might be explained by differences in management styles of dairy farmers (Roep et al., 1991; Leeuwis, 1993). It is also known that individuals rarely follow the same decision-making process when the same decision has to be made, due to a different cognitive style (Davis and Olson, 1985).

Bearing in mind that a subjective attitude and interest exist, an attempt will be made to make a general comparison of a few types (or methods) to come up with standards or reference values. These types are:

- HISTORICAL, the use of farm data from the previous year(s);
- PLANNING, the farmer's own planning or budgetting data;
- NORMATIVE, the result under normalised, conditioned production circumstances as the outcome of research to be used as reference;
- GROUP-COMPARISON, the average of a (selected) group of farms for comparison.

These four types are compared according to the requirements for standards as given in section 2.1. and as specified later in this section.

HISTORICAL

Historical data, or data from the previous year(s), are widely applied in accounts. A farm datum in the current year is compared with the value of the same variable in a previous year. This standard is no inconvenience for the organisation: there is no effort required in the development, the calculation is very easy and the comparison can be done instantly (no time interval). The effectiveness for management support is rather bad. Since the farmer is comparing his own results in the current year with his own results in one or more previous years, there is no insight in the quality of the management. The standard does furthermore not account for differences in e.g. weather or prices in successive years. Although this standard is implicitly farm-specific, its use is questionable when there are (structural) changes on the farm in the year of analysis, e.g. more milk production per cow or per hectare, more young stock. On the other hand, the farmer is interested in such an easy understandable standard, but it cannot be used as a goal or target.

PLANNING

Before the start of the year, the farmer can make a planning. He can set standards depending on his expectations, wishes and goals, but he cannot account for year effects because these are not known at the moment of the planning. Simulation models are often used as an aid. Planning standards are scarcely applied and it takes a lot of effort in the development. But they are readily available at the moment the account is finished.

The effectiveness is questionable, because there is no correction for external circumstances (e.g. the weather) in the year to come. Such circumstances can have such negative effects on the plan, that the value of these standards is doubtful. However, its value increases considerably when the planning is adapted during the period. But this is unusual.

Since the farmer develops the standards himself (or with an advisor), standards are corrected for farm-specific factors and for factors which the farmer assumes to occur in the coming year. When no extreme external circumstances occur, standards can be of much interest and serve very well as a goal. The understandability can decrease when knowledge about the applied assumptions and algorithms are not readily available at the time of analysis. The way in which the results are presented is important as well.

NORMATIVE

Some standards have been developed as a result of scientific research. These standards show what the results are under normalised production circumstances with good management and without incidental disturbing factors. The standards are instantaneously available and can be calculated with (mostly) a rather simple algorithm. This is favourable for the accountancy. However, other organisations (e.g. scientific research stations) require an enormous research effort for the development of such standards, although this will be done only once in a few years. The normative standards are useful for gaining insight in the management, but a drawback is that in general they do not account for year effects. However, corrections afterwards for e.g. bad weather may be done but is often neglected. Some standards account more or less for farm-specific factors, others do not. They are very powerful for setting a goal for the farmer, but when this goal is very far away, the standard might be so frustrating that the interest is questionable and fading away. It is not always clear to the user which factors are corrected for and under what circumstances these standards were developed.

GROUP COMPARISON

In the Netherlands standards that show the average performance or results of a (comparable/selected) group of farms in the same year are widely used. As with historical standards, there is no need for the development of algorithms. But for organisations that make use of them (e.g. accountancies), the calculation may be costly. Another drawback is that the calculation has to wait until all members of the group have their accounts finished. Since dairy farms are so diverse (see section 2.4), the standards will not be very farm-specific and it is therefore questionable if they can be used to get insight in management differences. However, disturbing year effects are limited. For its effectiveness, it must also be clear which (kind of) farms are in the comparison group.

Farmers want to compare their results with other, comparable, farms (De Hoop et al., 1988; Heinrich and Walter, 1989). This method of group comparison is very understandable since values are just averages. It is therefore not surprising that external standards are of much interest and widely used. But the understandability vanishes when the (kind of) farms are not known. When the external standards are just averages, they cannot be used as a goal. Presentation of the averages of a selected group of good farms (e.g. the best 25%) is only scarcely applied, maybe because many farms are required to perform the calculations.

In table 2.1, these four types are summarised and compared according to nine aspects. These aspects are the specifications of the necessary requirements of the standards (Section 2.1) and are important for judging if one or more types are of interest for application in the KBSs of the research at issue. The aspects are:

1. ORGANISATION, the burdening of an organisation (e.g. accountancy), which can be subdivided into:
   - DEVELOP, the effort of developing/adjusting the standard (algorithm);
   - CALC, the effort required to perform the calculation of the values;
   - TIME, the duration between finishing the account and the calculation.
2. EFFECTIVENESS, the effectiveness of the standard for management support, which can be subdivided into:
   - MANAG, the ability in giving insight in management differences;
   - YEAR, the degree of accounting for differences in years.
3. FARM&FARMER, account for the specific farm and the value for farmer, which can be subdivided into:
   - FARM, the degree in which the standard is farm-specific;
   - INTEREST, the interest of the farmer in having such a comparison figure;
   - GOAL, the way the standard can be used as a goal.
4. SIMPL, the simplicity of the standard to understand the meaning.

Table 2.1 Characteristics of four different methods for comparison: +++ = very favourable, --- = very unfavourable, / = or, ? = undetermined, etc (see text for explanation of terms)

	1.Organisation			2.Effectiveness		3.Farm&farmer			4.Simpl
	develop	calc	time	manag	year	farm	interest	goal	
HISTORICAL	+++	+++	+++	--	---	++/?	++	---	+++
PLANNING	--	+++	+++	+/?	-/+	+++	+++/?	+++/?	-/+
NORMATIVE	---/?	+	+++	+	--	?	?	+++	--
GROUP-COMP	+++	---	---	?	+++	--	+++	--/+	+++/?

None of these four methods actually meets all the requirements as set out in section 2.1 and are thus less qualified for application in the research described in this thesis. Standards based on historical data are, although very easy for the organisation to produce and for the farmer to understand, not effective for management support. Standards as a result of planning have doubtful effectiveness. The application of such standards in the research at issue is actually out of the question, since at the moment no accountancy uses them in their year-end account and it is not very likely that they will do so in near future. Normative standards have a drawback that it may be unclear what part of a deviation must be ascribed to year effects. Since farmers want to compare their results with others, normative standards might be less interesting and even frustrating for several farmers. This last drawback is of course absent with respect to averages of a (comparable/selected) group, the fourth method. Especially the disappointing comparative value, due to differences in farm structure, and the organisational burden, make this method unacceptable for this research. The limitation of such a standard is also remarked by Gekle (1990).

The method of FAS, a new method which will be explained in great detail in section 2.4, tends to be promising regarding the requirements from section 2.1. At the moment it is used by some farmers from the FADN network of LEI-DLO. In the end of this chapter, the FAS method will be evaluated with respect to the requirements in the same way as in table 2.1.

2.4 Farm-adjusted standards (FASs)

When data from many dairy farms are analysed, we see that - especially in this agricultural sector - there is a strong coherence between farm structural variables (like 'milk quota per hectare') and performance figures (like 'gross margin per cow'). It is therefore not so surprising that a dairy farmer wants to compare his results with *comparable* other dairy farms. Because of these structural differences, the problem always is to find a group of farms to compare with.

Suppose, a dairy farmer has data from no less than 625 farms at his disposal in the data base of his computer. He wants to compare his results with those farms that are comparable with respect to the following four variables: 'milk quota per hectare', 'number of young stock', 'amount of supplied nitrogen per hectare' and 'percentage of grassland from total area'. If we assume that these variables are independent and that there are five possible classes of values for each separate variable, then it can easily be calculated that on average $(1/5)^4 * 625 = 1$ farm, namely only his own farm, is in the selection.

With a selection coefficient of 1/3, which is rather rough, the farmer can compare his own results with data from about two dozens of farms

$((1/3)^4 * 625)$. But these farms may differ in milk yield, total area, breed, etc. It is not unlikely that a farmer will hide behind these last factors to gloss over bad results. A good or potentially good farmer will first try to find the cause of his problems on his own farm and look for ways of improvement, while a bad farmer will usually try to find the cause in factors outside his direct scope of decisions (Zachariasse, 1974). This behaviour will be denoted as the 'escape route' from now on.

FASs are developed to compare individual farm results with the 'average' of *comparable* farms, so that the above mentioned farmer's 'escape route' is limited. The 'average' is not calculated as a simple average of a group, but is rather the average after correction for several factors (see below).

FASs are developed and described in a publication by De Haan (1991). Since (1) FASs are essential in the methods and systems of DETECTOR (see chapters 6 and 7), and (2) it concerns a new method with good potentials, and (3) the method is not yet 1) described in English, the original description of De Haan (1991) is rewritten and summarised in the first part of section 2.4.1 2). The other sections deal with the use of FASs by farmers and organisations, and with an evaluation of this method regarding the requirements.

2.4.1 The method: the construction of the models

For the development of the arithmetical expressions, with which the FAS value can be calculated, the following multiple regression model (FAS model$_i$) is used:

$$y_i = c + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \ldots\ldots\ldots + \beta_k * x_{ki} + e_i \qquad (2.1)$$

where:

y_i = the i^{th} observation of dependent variable y;
x_{ki} = the i^{th} observation of the k^{th} independent variable x;
e_i = the error-term for the i^{th} observation of y with corresponding x values;
c = constant in the FAS model$_i$;
β_k = coefficient for x_{ki} in the FAS model$_i$.

---

1) In due time, the article "The development of farm-adjusted standards for the analysis of dairy farm performance, and its application in knowledge-based systems" by W.H.G.J. Hennen and T. de Haan will be submitted for publication.

2) T. de Haan gave permission for and reviewed the description of FAS in this section.

The constant and coefficients are estimated with the least-squares method of Genstat (Lane et al., 1988). There are regression analyses performed with returns, variable and fixed costs as dependent variables. Farm data that might influence those dependent variables are taken as independent variables. The relations that are estimated can be used as FAS models in the KBS for the calculation of FAS values for the various returns and costs just by filling in actual farm values for the independent variables. As a matter of fact, by using this methodology, the FAS values are corrected for the independent variables.

A suitable FAS model has to be found. This depends on the choice of the independent variables and how they are brought into the model (quadratic, interaction, etc). There has to be a theoretical relation between a dependent and independent variable.

From more than 300 specialised and representative Dutch dairy farms, stored data (FADN data base of LEI-DLO) are used for the estimation. This data set is extended with data from a few dozen other dairy farms. These farms are used for study purposes by LEI-DLO.

The regression analysis based on the empirical data is performed for each distinct year. The oldest FAS models stem from accounting year 1986/87. Models for the estimation of purchased feed and for the fixed costs are based on data from the five most recent years. Year effects are brought in by means of dummy year variables. These 'five-year' FAS models are updated each year. For all other dependent variables, only data from the most recent year are used. All FAS models are in fact year-dependent.

For some aspects, algorithms are also developed for the 25% highest and the 25% lowest performing farms with relation to that particular aspect. The choice of the percentage is in fact arbitrary, although Heinrich and Kalter (1989) remark that the use of quarters is generally adopted at the extension service in Germany. Gekle (1990) stresses the increasing popularity among farmers using quarters.

Let us illustrate the FAS method with an example from De Haan (1991), namely the FAS model for the dependent variable cattle credits in NLG per cow. In the model, three independent variables are included: 'the fat and protein corrected milk yield' (FPCM), 'the owned number of cattle per cow' (EGVEORMK), and 'the average breed of the herd' (BREED). The choice of these variables will be explained after this example has been presented.

The same model template is used for the FAS model of cattle credits based on data from all farms ($FAS_{cc\text{-}av}$), for the FAS model based on data from 25% of farms with the highest cattle credits ($FAS_{cc\text{-}hi}$), and for the FAS model regarding the lowest 25% of farms ($FAS_{cc\text{-}lo}$).

The three models are presented below.

$$FAS_{cc\text{-}av} = c_{av} + \beta_{1av} * FPCM + \beta_{2av} * EGVEORMK + \beta_{3av} * BREED \quad (2.2)$$
$$FAS_{cc\text{-}hi} = c_{hi} + \beta_{1hi} * FPCM + \beta_{2hi} * EGVEORMK + \beta_{3hi} * BREED \quad (2.3)$$
$$FAS_{cc\text{-}lo} = c_{lo} + \beta_{1lo} * FPCM + \beta_{2lo} * EGVEORMK + \beta_{3lo} * BREED \quad (2.4)$$

The development of the models $FAS_{cc\text{-}hi}$ and $FAS_{cc\text{-}lo}$ is actually done in two steps. First the realised cattle credits of all farms are matched against the outcome of the $FAS_{cc\text{-}av}$ model. Then all farms are sorted according to the deviation between the realised and the FAS value of cattle credits. The top 25% and the bottom 25% are selected from this sorted list. The second step is to develop a new model ($FAS_{cc\text{-}hi}$ or $FAS_{cc\text{-}lo}$), where only the selected group of farms (25%) is used as empirical data. Thus, values of all coefficients (the constants included) are different in the different models.

The selection procedure of the top 25% and the bottom 25% in the first step by De Haan (1991) assumes absence of heteroscedasticity or, in other words, absence of unequal variances of the error term e_i at different values of an independent variable. He could not find significant effects concerning the FAS models used in GLOBAL-DETECTOR (De Haan, 1990). Since heteroscedasticity might generally be expected (Theil, 1979), two equal deviations with FASs at two different values of an independent variable can have meanings that are *not* equal (Leneman, 1993). Accounting for heteroscedasticity is far from trivial, as can be concluded from the research done by Leneman (1993) on one FAS model. The uncorrected models (no accounting for heteroscedasticity) by De Haan (1991) will be incorporated in the KBSs of this thesis (Chapters 6 and 7), because (1) De Haan (1990) could not find significant effects, (2) the KBSs are meant to be global by giving just insight, and (3) maintenance and understandability is easier without correction.

Regression analysis starts with a specification of the regression model 2.2, which means that the independent variables (FPCM, EGVEORMK and BREED) have to be chosen from all possible variables (and combinations). The following requirements were taken into account by De Haan (1991) for the selection of variables:

- a theoretic logical relation between the dependent and the independent variable. Variables that make no sense to the user of the FAS model (farmer or advisor) are, despite of any possible significance, not very welcome to be included in the model;
- change of R^2 when a variable is included;
- the significance of the estimated coefficients by means of their t-values;
- as much uniformity as possible in the models of related dependent variables (e.g. the various components of purchased feed).

The three variables in the example are not only significant, but also logical for the user. Other FAS models for other aspects are also developed in accordance with these requirements. The reader is referred to De Haan (1991) for information on other FAS models.

FAS values can give the farmer insight in his position with respect to colleagues in a *comparable* situation, or those farmers with the same values for FPCM, etc. The farmer cannot hide behind the variables corrected for. The statement: '...my cattle credits are so low, because I do not have much young stock!', does not hold since the FAS value has been corrected for the independent variable EGVEORMK (number of cattle per cow). When this variable was not included into the FAS model, the farmer's reaction might be justified. However, with such a limited FAS model the unfavourable deviation could be 150 instead of 200. But it would be very wrong, if the farmer would ignore a deviation of 150 NLG totally, by using EGVEORMK as an excuse or as an 'escape route', and to act as if there is no problem at all.

To prevent the 'escape route', the independent variables in the FAS models have to be chosen with great care. In my opinion also non-significant independent variables must be eligible to be included into the FAS models to prevent the 'escape route', when these variables are often wrongly used as an excuse and when the impact after inclusion does not have an opposite sign as one would logically expect. This is not meant to provoke the statisticians or to mislead the farmer, but merely to protect especially those farmers who lack the knowledge or information for a balanced judgement or who lack self-criticism.

Of course, the 'escape route' cannot always be prevented. There might be numerous circumstances that are either too complex or cannot be described by the available data. This makes incorporation into the FAS model impossible. Remarks like '... I had bad luck!' or '... due to some investments, we had radical changes during the year' make interpretation of deviations tough.

FAS models must generally consist of independent variables whose values are more or less fixed or do not change in the short or medium term. They should not be effected by the operational or tactical management. The variable 'milk quota per hectare' is a good example of this. Other variables are disputable, like 'milk yield'. This variable can be influenced operationally by the variable 'amount of concentrates fed'. But this variable must be regarded as structural and important for correction, because this variable can easily be used as an 'escape route'.

Variables whose values are strongly effected by the management, e.g. 'amount of concentrates fed', must generally be avoided. But there may be exceptions, like 'nitrogen fertiliser in kg'. This (management) variable has much influence on the grass production and from that on the amount of roughage purchased. When the costs of feeding are not corrected for this variable, there is nearly no ground for comparison.

The total number of independent variables should be limited. The FAS model must not correct for everything, especially not for management factors (apart from exceptions). Simple models are in favour, they must be transferable to and understandable for the user. An advisor must know what variables are corrected for in the FAS model without looking it up in a manual.

It must be clear from the previous remarks that the development of FAS models is a far from easy and straightforward statistical exercise. The person in charge must possess not only statistical knowledge and experience, but he must also have knowledge and experience regarding dairy farm structure, dairy farm management, and farmer's behaviour. Only when great care is taken in the development of the FAS models and the use of the FAS values, FASs are powerful instruments, in combination with expert knowledge, for determining the farmer's position, for evaluating his management and for supporting management decisions.

2.4.2 The use of FASs

FASs can be used in the following ways:

I. FASs as standards for external farm comparison

The way in which FASs can be used as standards for external farm comparison is illustrated in table 2.2 and figure 2.1.

When the realised values for (corrected) milk yield, number of animals and breed for farm F in year Q are filled in in the FAS models (2.2, 2.3, 2.4) for year Q, the FAS values in table 2.2 are obtained. As shown, the realised value for cattle credits is about 200 NLG less than an average Dutch farm with the same milk yield, number of cattle per cow and breed has realised in that year. But the realised value for farm F is a bit higher than the average cattle credits of a group of *comparable* farms (25%) with low values for cattle credits.

Table 2.2 The different values of cattle credits per cow on farm F: realised, the farm-adjusted standard for the average of all farms, for the 25% of farms with the highest, and for 25% with the lowest value

	Realised	FAS average	FAS highest	FAS lowest
Cattle credits/cow (NLG)	581	788	1,007	551

II. FASs for giving insight in relations

With table 2.2, the farmer on farm F has an idea of his position with regard to cattle credits. This becomes even more clear in figure 2.1.

In this figure, the effect of varying values of the independent variable milk yield on the dependent variable cattle credits per cow is drawn. The other independent variables are kept constant on the farm's value for these variables. One of these other variables can also be varied and presented to the user in the same way while milk yield is kept constant. The small block in figure 2.1. denotes the farm's position of the realised value at a (corrected) milk yield of 7,755 kg. The line in the middle shows the FAS values based on data from all farms. The topmost line is based on data from the 'best' and the lowermost on data from the 'worst' 25% of farms.

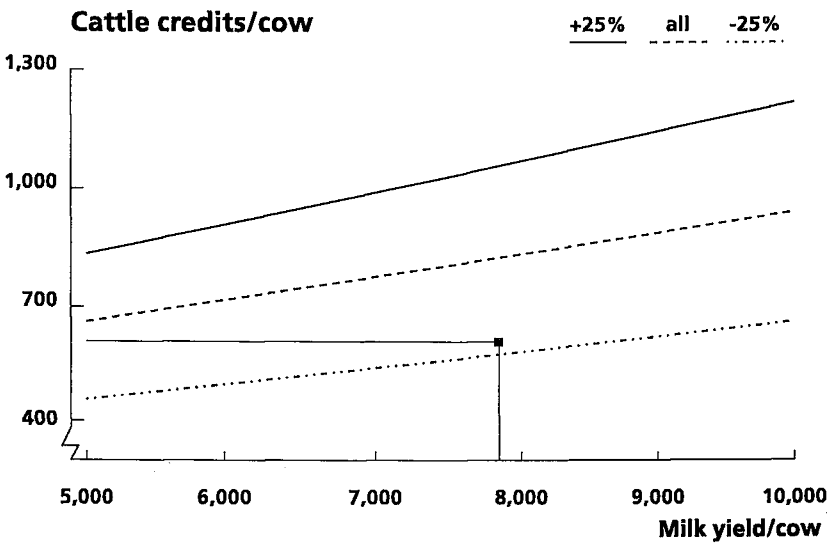


Figure 2.1 *Relation between the milk yield (X-axis) and the cattle credits per cow (Y-axis); and the position of a farm (little block). Output from GLOBAL-DETECTOR*

III. FASs for simulation (with farm-specific input-output relations)

Figure 2.1 can also be used to get an idea *WHAT* the expected cattle credits per cow would be *IF* the milk yield on farm F would have been 1,000 kg higher (8,755 kg), while the two other factors are kept constant (WHAT-IF question). When the little block is moved to the right, parallel to the lowermost line (-25%), the value is expected to be about thirty NLG higher. Here the *assumption* is made that the relation is the farm-specific input-output relation. However, the actual relation is not known from a farm, and it cannot be derived from the (accounting) data. An input-output relation is used instead where the curve depends on the position of the farm but is derived from the average of many farms. Nevertheless, this relation will be called farm-specific from now on.

The use as farm-specific input-output relation can also be illustrated in figure 2.2, where the effect of varying values of the independent variable nitrogen fertiliser (=input) on the dependent variable additional feeding costs per hectare (=output) is drawn. The costs for nitrogen fertiliser are included in the dependent variable.

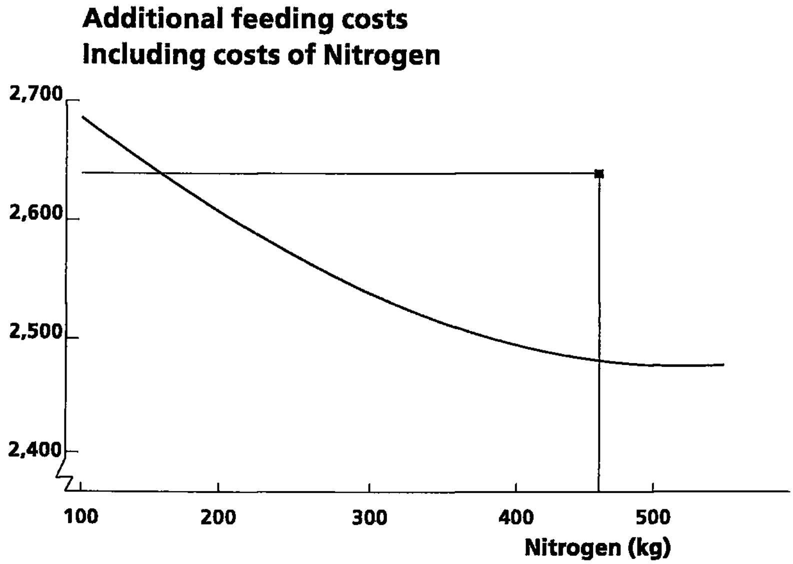


Figure 2.2 Relation between the amount of nitrogen fertiliser (X-axis) and the sum of additional feeding costs and costs for nitrogen fertiliser in NLG (Y-axis); and the position of a farm (little block). Output from GLOBAL-DETECTOR

The farmer must realise that the simulated result is just an approximation, as will be explained by means of figure 2.1. First, the line shows an average relation based on many farms, while the individual farm in question might show a different relation. Secondly, an increase of milk yield with 1,000 kg might go along with changes in farm structure and management on farm F. Although the two other independent variables are kept constant, the ceteris paribus principle might not hold due to possible other changes. Thirdly, the FASs calculated for farms with a milk yield of 8,755 (= 7,755 + 1,000) kg might in fact relate to different types of farms and farmers.

So the farmer must be careful in interpreting the lines like in figure 2.1 for WHAT-IF questions. Although the expected effects are only approximations, the farmer can get insight into the general effects quite easily. Since FAS models are already present, no additional models have to be developed for simple WHAT-IF questions when rough answers and indications are allright. In one of the KBSs of LEI-DLO (ENVIRONMENT-DETECTOR, see chapter 7), FAS models are incorporated in the arithmetical model. This model could be developed very rapidly by using FAS models.

IV. FASs as goals

FASs derived from empirical material are not meant as goals or targets. They are merely corrected averages. For example, according to the Dutch extension service the amount of fed concentrates are too high on most farms, resulting in high costs. The costs on an average Dutch farm will therefore also be too high, which means that the FAS based on the average farm may not be used as a target value. However, with this FAS value for the costs of concentrates, the farmer knows his position in relation to other farms. The FAS can be used as a goal by comparison with the 25% highest FAS for a return aspect or with the 25% lowest FAS for a cost aspect. With this information, the farmer also gets an idea of the relative magnitude of the deviation.

V. FASs for suggestions for improvement

Since deviations between realised and corresponding FAS values might contain useful information concerning the performed management on a farm, they are important information sources to form suggestions for improvement (advices). Such deviations can, with some other relevant data, be used together in a weighed way by the expert to make suggestions as to what actions the farmer can take to obtain a certain goal. Such use of FASs in the advice part (suggestions) of the KBSs will be explained later on (chapters 4, 6 and 7).

2.4.3 The use of FASs in organisations

Some accountancies supply the farmer not only with an account, but also with an overview of the results of other farms in the neighbourhood (group comparison). From this overview, the farmer gets an idea of his position. However, since each farm is somewhat unique, a good comparison is not always possible in this way. And there is another problem. The overview can only be supplied when all other accounts have been worked out as well. This may take some time.

The FAS, on the other hand, is a good figure for farm comparison and it can be supplied immediately with the account. The development of the FAS models may be based on data from the previous year. With additional price and quantity indexes, FAS values can be calculated immediately and sent together with the account. These indexes account only little for year effects (e.g. weather, prices), although sufficiently enough. Accountancies regard the timeliness as a great advantage of the FAS.

The provisional FAS values can be updated to definite values at the moment that all accounts are worked out. Using last year's FAS models as a template, the development of the updated FAS models can be done with the most recent (new) data.

2.4.4 Evaluation of FASs

Table 2.3 evaluates the FAS method according to the requirements specified in section 2.1 in the same way as the evaluation of other methods in section 2.3 and table 2.1.

Table 2.3 Characteristics of the farm-adjusted standard method: +++ = very favourable, --- = very unfavourable, / = or, ? = undetermined, etc (see section 2.3 for explanation of terms)

	1.Organisation			2.Effectiveness		3.Farm&farmer			4.Simpl
	develop	calc	time	manag	year	farm	interest	goal	
FAS	-	+	+++	++	++/0	+++	++	+++	-

Once the FAS algorithms have been developed, the yearly updates can be done in a short time. The calculation of FASs with these algorithms is rather simple with no delay in time. The effectiveness for management support and the accounting for farm and farmer are all strong aspects of this new method. However, when the FASs are just based on current indexes and on the FASs from the previous year, the accounting for year effects is little. But the accounting increases with the number of farms to be used for the actual calculation of the FASs.

A farmer can compare his results with good farms (25%) in the same situation, and this may be very stimulating because those farmers have actually proved it. A drawback is that the method has to be explained, and it must be clear to the user which factors are used for correction. The explanation facilities in the KBSs can cope with this drawback.

The FAS method is especially useful for the dairy sector which has a great diversity of farms and a strong coherence between farm structural variables and performance figures. This does not mean that the concept is of little or no value in other branches of agriculture. Compared to the performance of other methods (table 2.1), the FAS method has without doubt also potential for those branches where the number of farms is large enough. Some of these branches are characterised by different cultures and harvesting methods (horticulture) or different production systems (like housing in poultry and pig(let) production).

According to the requirements (section 2.1), it is my conclusion that the use of FAS as a standard is satisfactory enough for the research at issue. The method will be applied in the KBSs that will be described in some of the forthcoming chapters. At the moment it is not yet clear if FASs will turn out well in practice; there are no references known to me about the use of such a method. Our experience with farmers and organisations that worked with FASs or know its principles, tend to make us very optimistic for the use of FASs in the future. But it has to be proven.

3. KNOWLEDGE-BASED SYSTEMS FOR THE INTERPRETATION OF DATA FROM FARM ACCOUNTS

'Increasingly complex decision-making situations increase the probabality that wrong choices will be made. Knowledge-based systems can help increase the success rate, thereby improving farm performance.' (Wagner, 1993)

'The term 'expert system' is often abused by those impressed with the implications of the phrase. In reality, seldom does a system reach a level of competence that is deserving of the title 'expert'' (Evans et al., 1989)

3.1 Knowledge, knowledge transfer, and knowledge-based systems (KBSs)

The objective of analysing farm results is to find strong and weak aspects of the farm and its management. Those aspects are not simply the favourable or unfavourable deviations from standards. For example, since a high cost factor may result in a high return factor, that high cost does therefore not necessarily have to imply a weak aspect. Such deviations from standards, in combination with other factors and farm data, have to be judged carefully to make a sound diagnosis of the performance possible. The application of *knowledge* regarding the particular problem area is indispensable for such a judgement.

3.1.1 Knowledge

Humans use their knowledge to analyse farm results. Some persons perform this task better than others, mainly due to differences in experience. Generally, farmers only incidentally study their outcome. Although no one knows their farm better than they do, it is questionable if they have enough *experience* for this task. Narrowmindedness concerning their own farm may be a serious danger. Specialists or experts, for example advisors and extension workers, can see things on a farm or on an account in their true perspective. They are able to use their skills, based on a lot of training and experience, to perform a sound judgement and come up with some significant alternatives.

Human knowledge is complex. According to Lindsay and Norman (1977), 'Human knowledge is extremely extensive and everything seems to be related to everything else. Thus, when we describe human knowledge, the result can quickly appear to be complicated: the drawings look like cobwebs woven by aberrant spiders...'.

Humans use both declarative and procedural knowledge for a certain cognitive task. According to Gordon (1989), declarative knowledge consists of what people know about objects, events, static relationships between concepts, etc, and procedural knowledge is the knowledge about how to perform various cognitive activities. Procedural knowledge is 'compiled' knowledge after many years of experience, and therefore hard or even impossible to verbalise. Experts are sometimes even unaware of what they know.

Only a limited number of farmers makes use of the knowledge from advisors, etc, for the evaluation of their year-end results, although farmers have generally difficulties in analysing farm data on their own (chapter 1). To have widespread access to the knowledge of the best experts there are, distribution by means of KBSs may be of interest. The knowledge has to be extracted from advisors and experts and transformed to such a system.

3.1.2 Knowledge transfer

The process of extracting the knowledge from an expert (elicitation), and formalising and coding it into a computer programme, is called knowledge acquisition or knowledge transfer (Hayes-Roth et al., 1983). Knowledge acquisition is such a difficult and time consuming phase in the development of a KBS, that it is even called 'The bottleneck' by many Artificial Intelligence scientists (e.g. Feigenbaum, 1977). The problems originate mainly from the 'compiled' procedural kind of knowledge that cannot be verbalised. Johnson (1983) talks about 'the paradox of expertise', referring to the fact that the more persons know, the less they are aware of what they know. 'The knowledge we most want to represent in a KBS, often turns out to be the knowledge the expert is least able to talk about'.

A great number of different methods exists to extract knowledge. The choice of which method or methods to use mainly depends on the domain, the task of the expert, the expert himself, and the acquaintance of the knowledge engineer (the builder of the KBS, see below) with these methods. The reader is referred to the extensive literature on this subject. For example, in 1991 I described and compared a dozen techniques for knowledge elicitation with special reference to the agricultural economics domains (Hennen, 1991).

Procedural knowledge is the type most used by the expert, but the knowledge that is gained by nearly all acquisition methods and tools is declarative (Gordon, 1989). This means that not the true knowledge of an expert can be captured, but merely a meagre extract of it. Inferences done with a KBS that contains and uses knowledge to perform at high levels, is also far from the real decision-making process performed by human experts. Breuker and Wielinga (1989) stated that experts understand the problem and see or refine the solution, while a system only solves problems by reasoning as a well informed and systematic novice would do, because the process of understanding and seeing appears to be inaccessible. In this perspective, the very fashionable term *expert system* will not be used in this thesis. It might be misleading and it will not meet the expectation that the proposed systems will perform like a human expert. According to Waterman (1986), all expert systems are in fact KBSs, while the converse may not always be true.

3.1.3 Knowledge-based systems (KBSs)

KBSs originate from research in Artificial Intelligence. These computer programmes contain explicit domain-specific knowledge in a very narrow domain and are able to use that knowledge to solve problems, to make or support decisions and to generate conclusions from several data with respect to a certain case or situation. Knowledge may stem from different sources, although predominantly human (expert) knowledge is used in most systems. In many real situations rules of thumb are used by the expert to infer conclusions from the information he has gained. It is therefore not so surprising that the majority of KBSs is of a rule-based type, in which so-called IF..THEN statements represent such rules of thumb.

Figure 3.1 shows an example of a rule expressed in an IF..THEN statement. This is a modified rule from the KBS GLOBAL-DETECTOR (chapter 6).

IF:	P1:	milk quota per hectare is not very_high AND
	P2:	milk yield is at least rather_high AND
	P3:	amount of purchased feed is lower than on comparable farms AND
	P4:	cattle credits minus animal costs is quite_low compared to others
THEN:	R1:	"Try to increase earnings by selling (more) breeding cattle"

Figure 3.1 Example of an IF..THEN statement in rule-based systems

Such a rule represents the deduction that can be made from facts in the IF-part and the corresponding data of a particular case (i.e. a farm). When all four facts in the IF-part match the data so that all premises or conditions are true, the conclusion in the THEN-part is also true. The THEN-part may also be an action or it directs the control of the programme. An inferred conclusion may thereupon be used by a condition from another rule from the KBS. The conclusion is not inferred when one condition is not met, regardless of the other conditions. Thus compensation is not possible. In section 3.2, we shall see that the inference is not so strict when uncertainty is applied.

The symbolic representation of the domain knowledge (i.e. the facts and rules) is contained in the *knowledge base* of the KBS, actually the most important part of the system. In the knowledge base, rules are added, deleted and modified in a declarative programming style. The contents of the knowledge base is mostly written in pseudo-natural language, e.g. like the rule in figure 3.1. Since the programmer has to worry less about the control strategy or general problem-solving knowledge, the programming of the knowledge base is quite easy, making a rapid development of prototypes possible. A modification in one part of the system has not so much impact on other parts of the programme, in contrast to programming 'conventional' systems. This separation between the domain knowledge in the knowledge base on the one hand, and the control structure for utilising that knowledge on the other, is in fact the most important difference between KBS and 'conventional' systems.

The control structure contains the general problem-solving knowledge, and it is called the *inference engine* (Waterman, 1986). This part of a KBS is the means to use the knowledge in the knowledge base effectively. As a matter of fact, the facts and rules are processed by the engine's method or methods to yield conclusions or actions. Two commonly used methods used in inference engines are forward (data driven) and backward (goal driven) reasoning. While the first method tries to match the facts and the case data to infer conclusions or new facts, the backward method tries to prove the concept in the THEN-part as if it would be a hypothesis. It depends heavily on the kind of domain and the objective of the KBS which (combination of) method(s) of the inference engine is most suitable.

The user of a system must have confidence in the results, so he has to understand the reasoning process of the KBS and the justifications of the conclusions. The *explanation facilities*, which are part of the *user interface* of a KBS, provide that information. They show how a certain conclusion is reached, normally describing the rule or a part of the sequence of rules that led to the conclusion (Waterman, 1986). Such a kind of explanation facility is often criticised, they are 'flat' and often not clear to the user. In chapter 4, additional comments on these facilities are made.

The person who acquires the knowledge from the expert and who builds the KBS is generally one and the same person, and is called the *knowledge engineer*. Nowadays very sophisticated building tools for KBS are available on the market. They speed up the development considerably. Such tools have built-in inference engines, so that the knowledge engineer need not build the inference engine for a new application. The knowledge engineer's task is mainly reduced to the acquisition of knowledge and the development of the knowledge base. Recent developments in knowledge acquisition tools facilitate the development of KBS even more, and decrease the importance of the knowledge engineer's role.

KBS building tools have to be chosen very carefully. Preferably the choice is made when it is clear how to represent the knowledge and what the requirements concerning hardware and user interface are. Sometimes, suitable KBS building tools are unavailable or not flexible enough to meet all the requirements. In such cases, systems may be developed from scratch with programming languages, e.g. the Artificial Intelligence's languages Prolog and Lisp. A KBS development with such a language does not only require much effort to learn, but will also be very costly and time consuming, since the control structure and the user interface have to be programmed.

KBSs for agriculture are generally built with the aid of tools, only very few examples exist where Prolog or LISP is used (e.g. Evans et al., 1989). Due to the required flexibility, all the KBSs from LEI-DLO are developed from scratch in language muLISP (Soft Warehouse). This language, a dialect of the standard Common Lisp (Steele, 1984), runs on a PC, is relatively fast and consumes only a very small amount of memory. The reasons for the development from scratch of the KBS Cattle Credits-DETECTOR, a system for the analysis of cattle credits on dairy farms (Hennen, 1989), have been the required speed and internal memory capacity. One reason for the development of GLOBAL-DETECTOR (chapter 6) and ENVIRONMENT-DETECTOR (chapter 7) from scratch is the need to incorporate the newly developed inference methods IMAGINE (chapter 4) and FUZZY-DETECTOR (chapter 5). Other reasons are the enormous flexibility and the requirement that both systems must perform other tasks as well, e.g. calculations and graphical presentations. Such combined systems may therefore be called hybrid systems.

It is important that a KBS performs satisfactorily enough for what it is intended for, and that it can be a good alternative for a human expert. A KBS can even be advantageous compared to the expert, due to the accessibility, speed, consistency (i.e. more structured), a greater availability of data and information, and so on. The combination or integration of computer power for calculations (e.g. an arithmetical model) and a KBS for reasoning may even be likely to perform better than a human. Especially such combined or hybrid systems will be of particular interest for agriculture (Barrett et al., 1985; Stone, 1989; Harrison, 1991; Wagner, 1993).

3.2 The management of uncertainty in KBSs

Knowledge for a KBS does in general not have a logical structure, where conclusions are inferred from several facts in a logical way, like the knowledge contained in the Flora (Swaan Arons and Van Lith, 1984). In fact, conditions and inference rules may be uncertain and data are sometimes imprecise, unreliable or even missing. These problems that have to be managed in a KBS will be denoted by the term *management of uncertainty* from now on.

There are different methods to manage uncertainty. The well-known Bayesian probability theory will not be discussed here, since most KBS domains do not meet the requirements 1) for its application (Luger and Stubblefield, 1989).

This section will only briefly go into the heuristic approach from the certainty theory as used in the MYCIN system (Buchanan and Shortliffe, 1984), into the management of uncertainty using fuzzy sets and into the uncertainty concerning the judgement of numerical data with 'numerical knowledge'. The heuristic approach is often applied in KBS when uncertainty is at stake, while fuzzy sets may be of increasing importance in the future.

3.2.1 Heuristic approach for the management of uncertainty

To explain the heuristic approach, we take an abbreviated form of the example of figure 3.1:

IF: P1 and P2 and P3 and P4 THEN: R1(0.7)

If all premises or conditions P are certain and absolutely true, then the conclusion R1 is true with the certainty factor (CF) of the IF..THEN rule: 0.7. This CF represents the expert's confidence in the conclusion, which might be a real number in the interval [-1,1], or ranging from absolutely false to absolutely true.

Premises are not always absolutely true. It may be ambiguous to give the truth content for 'P1:milk quota per hectare is not very_high' regarding a particular case, since an individual does not know exactly what is meant by 'very_high' and how a particular case datum matches with 'very_high'. Suppose one answers a question concerning premise P1 with 'probably true'. If the CF for this statement is 0.8, then the confidence in the conclusion R1 will certainly be lower than 0.7.

---

1) All statistical data on the relationships of the evidence with the various hypotheses must be known, all relationships between evidence and hypothesis must be independent and continuous updates of statistical data are needed during consultation (Luger and Stubblefield, 1989).

Formal algebra exists to perform calculations with CFs. The combined CF of all premises is the minimum of the individual premises when the connective is 'AND'. When, for example, the CFs for P1, P2, P3 and P4 are 0.8, 1, 0.4, 0.6 respectively, the combined CF is 0.4 (the minimum). With an 'OR' connective, it would be the maximum. The truth content of the conclusion will be the product of the combined CF (0.4, or the minimum of 0.8, 1, 0.4 and 0.6) and the CF of the rule (0.7): 0.4 * 0.7 = 0.28.

A drawback of taking the minimum is that it concentrates only on the worst premise. It does not matter how good or bad the other premises are, as long as their CF is higher. It is intuitively wrong that the CFs 1, 1, 0.4 and 1 are treated exactly the same as 0.4, 0.4, 0.4 and 0.4, especially when the importance of the worst aspect is not higher than that of others. For a sound judgement, a sort of compensation must generally be possible.

3.2.2 Fuzzy sets for the management of uncertainty

In section 3.2.1, it was already mentioned that the meaning of the linguistic term 'very_high' from the premise 'P1:milk quota per hectare is not very_high', is ambiguous due to its vagueness or inexactness. To use such linguistic variables 1), it seems to be inappropriate to apply precise quantitative analysis with the tools of statistics or with mathematical terms (Jain, 1977; Zadeh, 1973). However, a linguistic treatment seems essential to build KBSs that have a 'humanly-perceived approach' (Freksa, 1982), especially since the knowledge from a human expert is usually derived in linguistic terms (Negoita, 1985). Humans manipulate fuzzy concepts and respond to fuzzy instructions (Bellman and Zadeh, 1970). The key elements in human thinking are not numbers, but labels of fuzzy sets, that is, 'classes of objects in which the transition from membership to non-membership is gradual rather than abrupt' (Zadeh, 1973).

The theory of fuzzy sets suits the way humans think (Zadeh, 1973) and it offers a rather appealing way to incorporate subjective evaluations in knowledge bases. Zadeh (1965) laid the foundations for the fuzzy set theory for measuring vagueness by possibilities and thereby measuring the *meaning* of information.

Zimmermann (1991) notes that the reasons for the application of fuzzy sets in KBSs are (1) more 'natural' communication, (2) the imprecise nature of human knowledge, and (3) need to deal with (management of) uncertainty.

---

1) A linguistic variable represents the fundamentally imprecise human perception of physical reality, and it differs from a numerical variable in that its values are not numbers but qualifying words as good, profitable, and so on. The linguistic value for an attribute can then be expressed as numeric through the use of fuzzy sets (Negoita, 1985).

For a theoretical background of the application of fuzzy sets in KBS, the reader is referred to Negoita (1985) and Zimmermann (1987). Graham (1991) gave a short history of the attemps for this application and surveys the use in commercial KBSs. Zimmermann (1987) and Zimmermann (1991) described a few applications more in detail.

The CFs from section 3.2.1 qualify the certainty of diagnosis by indicating the heuristic strength of the rules. With fuzzy sets, on the other hand, this can be indicated by the relevance of the case data to the rules (Coughlan and Running, 1989). When, for example, a case datum is expressed in the linguistic and vague term 'above_average', this can be matched with the vague expert's condition 'very_high' of P1 from the rule of our example. The method FUZZY-DETECTOR has been developed to be able to cope with such uncertainty in KBSs based on rules. This method as well as some concepts of the fuzzy set theory are described in chapter 5.

The theory of fuzzy sets is very suitable for the management of uncertainty in KBSs, since it provides a systematic basis for and inferring from and representing imprecise rather than precise knowledge by fuzzy mathematics (Zadeh, 1983; Gaines et al., 1984). Especially in farm management there are many uncontrollable factors which economists may view as the risk and/or uncertainty problem of the farm business. Nagaki (1992) notes that a probable next step in the software development for aiding farm management in decision-making will be the introduction of fuzzy set theory. Nagaki is also very optimistic in expecting that the application of the fuzzy set theory will help to promote a widespread on-farm computer usage in the future. However, Graham (1991) states that its use for modelling uncertainty in KBS is very controversial. FUZZY-DETECTOR (chapter 5) can therefore be used to explore the usefulness of fuzzy sets for the management of uncertainty, especially regarding the evaluation of farm data.

3.2.3 Uncertainty regarding the knowledge of weight and importance

In most cases farm data are not imprecise, but have rather strict values. A case datum for quota per hectare, for example, might be 12,584 kg. When the corresponding condition is 'P1:milk quota per hectare is not very_high', the farm datum can be matched by FUZZY-DETECTOR in the same way as shown in section 3.2.2 (see chapter 5).

Domains where data have to be analysed, e.g. in financial or economic domains, are mostly characterised by a quantitative nature. Conditions or premises of the rules in the rule base of a KBS may then look like 'P1:milk quota per hectare < 20,000', instead of the above mentioned expression from the example in figure 3.1. Farm data are also numerical, e.g. 12,584 kg. Since 12,585 < 20,000 is true, the first condition is also true. However, it is doubtful that a farm datum of 19,990 has the same meaning or cognitive weight regarding the conclusion as the value 12,584 has, although 19,990 < 20,000 is also absolutely true. It is furthermore doubtful that 19,990 and 20,010 are treated totally different, while both 12,584 and 19,990 have the same weight. It is intuitively clear that a farm datum must have a certain increasing or decreasing weight with regard to the conclusion, and not just an abrupt boundary like 20,000. Such a weight plays a role in the judgement by the expert.

Another aspect is that different concepts may be of different importance for the conclusion. When we take the third (modified) premise from our example in figure 3.1, 'P3:amount of purchased feed < amount on comparable farms', it does not need much imagination to conclude that a difference of 1,000 for this premise and a difference of 1,000 for the first one (P1) have to be treated distinctively with regard to the conclusion due to difference in importance.

Weight of the values and relative importance of the various conditions is implicit expert knowledge. We have to face imprecision and uncertainty regarding these aspects. In the next chapter, the method IMAGINE will be presented which tries to approximate expert's knowledge regarding this kind of uncertainty.

3.3 Application of KBSs for analysis of year-end results

According to the considerable number of reported applications, KBSs seem especially useful for the interpretation of data sets. The usefulness of KBSs for the analysis of year-end results of dairy farms has already been motivated at length in the first chapter of the thesis. Harsh (1988) advocated the use of KBSs for the analysis and interpretation of the accounts from an accounting record system to provide the manager with valuable information 'in much the same fashion as a knowledgeable or experienced financial document analyst'.

In this section, only a few examples of KBSs will be reviewed. This is followed by a motivation for the development of the DETECTOR 1) methods and systems, that are described in the following chapters.

---

1) The term DETECTOR (Discursive Expert for the Technical and Economic Control, Testing and Opinion-formation from Reports) is used as a suffix for a number of (related) research products from LEI-DLO regarding Artificial Intelligence and KBSs (e.g. Cattle Credits-, GLOBAL-, ENVIRONMENT- and FUZZY-DETECTOR). The method IMAGINE (chapter 4) is also a research product from LEI-DLO.

3.3.1 A short review from the literature

Reported KBSs, which analyse year-end results, address predominantly the financial condition of the farm. According to McGrann et al. (1989), '.. expert system technology will be a valuable tool to enhance knowledge delivered by agricultural economics. Economics and finance are areas where expertise is often limited, leading to inadequate use of data and analysis tools by producers, lenders and educators. Expert systems offer a significant delivery technology.' Bouwman (1982) saw the analysis of financial statements as a diagnostic process: the analyst examines the data from a firm (the patient), and is looking for clues (symptoms) to try to locate possible problems (diseases).

Phillips and Harsh (1987) developed a prototype KBS for the analysis of dairy farm financial records, based on the data from income statements and balance sheets. Only a few data were used to come to conclusions that predominantly concern the financial situation. The knowledge base was structured in a decision tree, without using CFs. This caused problems when farms were near the borderlines between different categories of firm position. The proposed solution of bringing in additional sets of rules, appears to be a bit ad hoc (see section 3.3.2), although the adjusted prototype performed satisfactorily.

An often cited KBS is FinARS (Boggess et al., 1989), which provides from a very small data set a 'quick and easy' evaluation of the financial health of a farm business. This KBS is quite similar to the system developed by Phillips and Harsh. FinARS not only provides an initial financial interpretation, it also diagnoses potential problems and gives suggestions for improvement. McGrann et al. (1989) describe the Agricultural Financial Analysis Expert System (AFAES), which includes software to make summaries and graphic presentations of the analysis and a diagnostic analysis of the financial statement data. Dobbins and King (1988) present a KBS to assist managers in interpreting year-end farm bussiness summaries from crop-hog farms. The KBS FARMEXPERT, which was reported very extensively by Wagner (1992), analyses the profitability of farms by comparing the farm in question with the averages of other farms. The system gives hints to improve the situation of farms in a poor position. Fillatre and Moreau (1991) reported COFINE (COmmentaires FINanciers de l'Entreprise). This KBS helps accountants to make a financial analysis of the farm and can support strategic planning. Huirne (1990) developed the personal computer system CHESS, that analyses the economic and technical records of individual swine breeding herds. This system tries to find strengths and weaknesses by combining and evaluating deviations between performance and standards.

In the Netherlands, only recently accountancies are developing systems for the interpretation of financial data. Although they are no KBSs, some are nevertheless worth mentioning. KASA (Breembroek, 1991) detects the main problems and strong items concerning the management of individual arable farms by comparing individual farms with the average and best 30% of farms in the region. For farms with pigs, a comparable system is under development (Bottleneck Analysis System Pig husbandry, BASP or KASVA; Baltussen et al., 1993b).

Several KBSs are in use at banks for granting loans based on financial results from farms (Carrascal and Pau, 1992). There are also many KBSs that support operational decisions. These systems, which will not be discussed here, are mainly based on more or less technical data that generally do not stem from the year-end account.

3.3.2 Motivation for the development of DETECTOR

Many KBSs are developed for the analysis of the financial data from farms (section 3.3.1). It is rather disappointing that most systems are not used in the way expected or not used at all. It goes beyond this study to make an investigation of the reasons for that.

Without ever claiming that the DETECTOR concepts, methods and systems as described in this thesis will be successful, and without suggesting that the used concepts and methods and described systems reviewed in section 3.3.1 are less appropriate, the motivation of my approach shall be outlined. My approach is different from the other approaches with respect to many of the following aspects:

1. Much emphasis will be placed on the requirements that dairy farmers themselves have expressed in a study by De Hoop et al. (1988). These and some other requirements have been described in the list of requirements in section 1.4. Many reported KBSs in section 3.3.1 fail to meet these requirements.
2. One of the farmers' requirements is that the system should be aimed at tactical decision-making by analysing the aspects of gross margin (De Hoop et al., 1988). Dairy farmer's concern are the returns and variable costs, while long-term decisions regarding fixed costs only occur incidentally. Dairy farmers are primarily focussed on efficient production, expressed in the gross margin from the year-end account (Zachariasse, 1990). Zachariasse observed that decisions concerning production are not only numerous, but they are also difficult to transfer to others. This is due to the complex circumstances in which these decisions take place. Because production is closer to the daily interest of the farmer and has an important impact on the financial results, the analysis and diagnosis of the *financial condition* of the farm (likewise most of the mentioned KBS in section 3.3.1) will *not* be dealt with in this thesis.
3. Potential users (e.g. dairy farmers) must be involved in the design and development of the methods and prototypes as soon as possible. Wain et al. (1988) noted that users are better capable of criticising an existing system than specifying or anticipating their requirements without a system. A prototypical development is often

advocated for the development of agricultural KBS (e.g. Berry et al. 1991; Gordon et al., 1987; Phillips and Harsh, 1987; Wain et al., 1988; De Hoop, 1991).
Advantages are:
- ability to represent a common reference point for discussion;
- motivation of the persons involved in the development of the KBS;
- prevention of difficult changes of the KBS later on;
- account for farmer's information need and decision behaviour and his wishes and requirements;
- better ensurance of user's acceptance of the KBS;
- earlier insight in perspectives for the market for the KBS;
- rapid illustration of the basic forms and functions of the methods and systems for purposes of demonstration and presentation of (scientific) results;
- better results when the experts can react.

4. The proposed systems must not give the ultimate answer, but must be used as an aid and should give insight in (a) the aspects that are important for decision-making, (b) the relevant inference processes, and (c) the direction of the resulting outcome (De Hoop et al., 1988). Generally, KBSs suit such an approach when the use of the explanation facilities are emphasised. The proposed systems must also be able to present results graphically in a flexible way.
5. As already indicated in chapter 1, a top-down approach will be advocated, so that the analysis is performed *globally* from a few data from the year-end account only. The strong and weak aspects found and the suggestions are therefore more or less plausible instead of totally true or false. The provided explanations and justifications with the explanation facilities will decrease user scepticism (Evans et al., 1989).
6. The fact that conclusions are more or less plausible due to the global approach means that the management of uncertainty is important. Uncertainty has already come up for discussion earlier in this chapter. In chapter 5 the method and tool FUZZY-DETECTOR deals with the management of uncertainty. Nearly all KBSs described in section 3.3.1 lack methods for the management of uncertainty.
7. Farm results must be compared with the results from comparable other farms as best one can, to present a reliable image of the farm's position and to counter the escape route (chapter 2, farm-adjusted standards). Only in a few existing KBSs, farm comparision is performed, and mostly with the average of a group of farms, without accounting for farm-specific structure and circumstances.
8. The knowledge bases of some of the KBSs mentioned in section 3.3.1. are structured in a decision tree. It is surprisingly, however, that the problem with the borderline values is not addressed except in one KBS. But it appears that provisions for borderline values (with additional sets of rules) done by Phillips and Harsh

(1987) are a bit ad hoc and might be an extra claim on the system itself and on the time for development and maintenance of the system, and hence, a possible loss in understandability. In DETECTOR, methods will be developed to tackle problems with borderline values (see section 3.2.3., and the chapters 4 and 5).

9. Development of a KBS is generally cumbersome and may cost a lot of time and money, especially due to the tedious knowledge acquisition. The approach of DETECTOR must thus not only be oriented towards the development of KBSs, but also find ways and methods to speed it up (chapters 4 and 5). Such methods should also ease maintenance, an issue not often fully addressed as remarked by Harsh (1991).
10. KBSs have to be developed in a flexible environment, without being constrained too much from software and hardware. Existing 'shells' are mostly too rigid and do not meet the requirements for building the tools for the methods and the KBSs as described in the following chapters. LEI-DLO's choice for the development 'from scratch' has been motivated in 3.1.3.
11. The KBS to be developed must be usable for individual consultation on PC by different groups of users. There should be a farmer's version and a coach version (for advisors) of the same system. It must also be possible to use the same system for automatically producing results on paper for a great number of farms, so that accountancies can add them to the original account with a minimum of time and costs.
12. Most reported researches in section 3.3.1 were oriented towards a final product: a KBS. In the DETECTOR approach, the KBSs are not the only products. Of special interest are the methods and tools for their use in other research and development activities (see e.g. chapter 8).

These motivations and implicit requirements guided the development of the methods and tools (chapters 4 and 5) and KBSs (chapters 6 and 7) of DETECTOR.

4. IMAGINE: METHOD FOR ACQUISITION, REPRESENTATION AND PRESENTATION OF KNOWLEDGE IN QUANTITATIVE DOMAINS

*'.... the expert is asked to **imagine** how he would solve (..) problems.'* (Breuker and Wielinga, 1984)

4.1 Introduction

The quantitative character of the data hampers the development of knowledge-based systems (KBSs) in technical, economic, or financial domains. Such data are continuous and may therefore attain numerous different values. Tools for building traditional rule-based systems 1) are rooted in a two-valued logic, and thus the rules must be executed in an all-or-nothing manner (Whalen and Scott, 1983). They cannot deal with continuous variables.

The average milk yield on dairy farms, for example, can have a value ranging from less than 5,000 to more than 10,000 kg. A subdivision in several classes, when continuous variables are made discrete, seems necessary to use them in traditional rule-based systems. When a condition says that 'milk yield > 7,250', the values 7,255 and 10,000 are treated the same way but the values 7,255 and 7,245 are treated differently. At a value of 7,245, the resulting conclusion will never be true, though other conditions might be strongly supporting the conclusion. The expert shall certainly disagree in such cases. By making ten distinct classes, in each class an interval of 500 kg milk, the condition would be 'milk yield > 7,000 and < 7,500'. When other conditions are treated in the same way, the knowledge regarding the conclusion at hand is much better modelled, though nothing can be said about the magnitude of the relevance of the conclusion.

However, when all situations are covered exhaustively in a reliable way, the result is an unmanageable magnitude because of combinatorial explosion. If, for example, there are n different input data required for a given conclusion and the value of datum i can be member of K_i (i=1,..,n) distinct classes, then the *theoretical* maximum number of rules to cover all situations for that particular conclusion is $K_1*K_2*...*K_n$. The desired

1) Most knowledge-based systems are rule-based, where the knowledge is expressed as IF-THEN statements. In the text they will be called traditional, since the proposed methods in this and the next chapter have a different approach.

reliability of the outcome is determined by the number K_i of classes for each datum. This level of detail in a chunk of knowledge is called granularity (Waterman, 1986).

The development of a knowledge base in such a way with many related rules causes serious trouble regarding the knowledge acquisition, the implementation (takes up too much computer memory), the clearness of presentation, the explanation facilities and the maintenance of the knowledge base. Exhaustion is therefore impossible in most quantitative domains. In 1991 I described and compared a dozen techniques for knowledge elicitation (Hennen, 1991). This was done with special reference to agricultural economic domains. Because none of these techniques tackled the problem of combinatorial explosion, they will not be presented or discussed here.

In an earlier attempt by LEI-DLO to build a KBS in an economic domain, this problem of combinatorial explosion occurred. After half a dozen farms had been analysed by the expert, it became intuitively clear that building a reliable KBS in reasonable time seemed impossible in this way. The expert's task for this attempt was to analyse actual accounting records and to make conclusions regarding strong and weak aspects of farm management. For each farm, he used a finite data set from the record and derived a finite set of possible conclusions. The expert's knowledge consisted of deciding which data to use for a given conclusion and determining how the value of each datum had to contribute to the relevance of the conclusion. The elicitation of this kind of knowledge was aided by the method protocol analysis (Ericsson and Simon, 1984).

To face this problem in the right way calls for a structural approach, far more different than the attempt by LEI-DLO described above. The structural approach to the problem of combinatorial explosion will be the application of the method IMAGINE 1), with *fuzzy boundaries* and *compensatory mechanisms* as the main characteristics. Additional advantages of IMAGINE are the facilitation of knowledge acquisition and comprehensive explanation facilities (in contrast with rule-based systems). With this method, cognitive models of the expert regarding conclusions can be developed. Such models can be used in KBSs, eventually with explicit arithmetical models (e.g. simulation models). Only cognitive models are used in the systems of DETECTOR (Chapter 6 and 7), because the expert was able to infer conclusions without the aid of additional arithmetical models.

---

1) Introspective Method to Acquire Goal-directedly Indefinite Numerical Expertise.

The objective of this chapter is to explain the method IMAGINE through an example. Not only the algorithm of IMAGINE will be described. The roles of the expert, the knowledge engineer and the user - who all make use of this new method in a different way - are also brought up. This is followed by a comparison with traditional rule-based systems regarding explanation facilities. In the discussion, the prospects and limitations of IMAGINE are outlined.

4.2 Model-based knowledge acquisition with IMAGINE

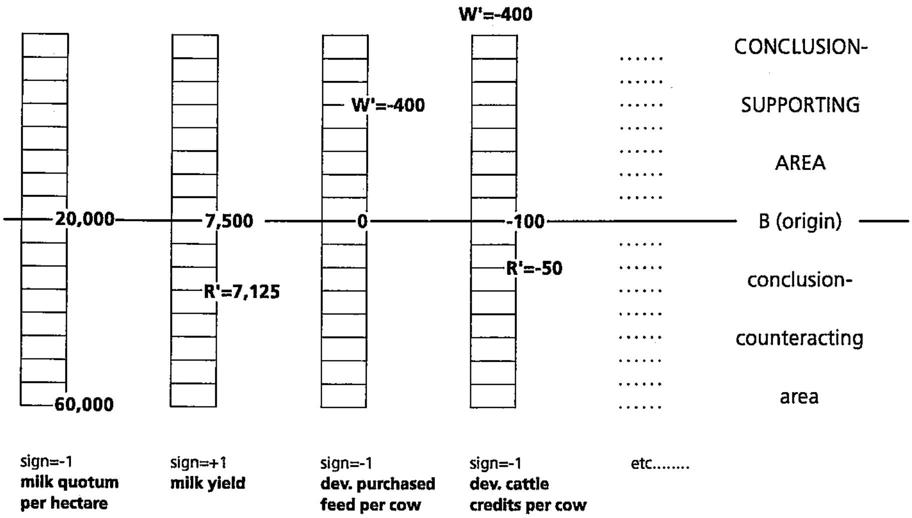
IMAGINE is a tool to facilitate knowledge acquisition for quantitative domains. The most difficult and time consuming process in building KBSs is knowledge acquisition: the transfer and transformation of problem-solving expertise from some knowledge source to a programme (Buchanan and Shortliffe, 1984). The process is often viewed as the 'The Bottleneck' in the literature.

The knowledge engineer first spends some time with the expert to explain knowledge acquisition by IMAGINE and how conclusions are inferred by the functions of the algorithm (see section 4.3). After the method is made clear, the expert can formulate his knowledge without the intervention of the knowledge engineer. Because of this independence, the expert can perform this task when and wherever he likes without being influenced or guided by the knowledge engineer.

The expert has to bring to mind and to write down all conclusions which are of interest for the domain. Each individual conclusion is then put through knowledge acquisition by IMAGINE. One of the conclusions from the domain of GLOBAL-DETECTOR (chapter 6), the suggestion 'Try to increase earnings by selling (more) breeding cattle' (Conclusion_8), will be used as an example in this chapter. During knowledge acquisition, the expert is asked to concentrate on this conclusion, which may be of importance on a particular farm. After he has formed a picture of this in his mind (= **imagine**), he is asked to write down essential information about this conclusion on a standard form (figure 4.1). This introspective process is here called *'backward knowledge acquisition'*.

The expert has to write down on the standard form (like in figure 4.1) all the information needed to infer the given conclusion 'Try to increase earnings by selling (more) breeding cattle', i.e. the name of the conclusion, the input data needed (variables), and the parameters (numerical values) required for the functions of the algorithm described in section 4.3. All information that stems from the expert, is marked as bold characters in this figure. The information is sufficient for the tool of IMAGINE to infer the relevance of the conclusion of figure 4.1 for any farm.

NAME OF CONCLUSION: **"Try to increase earnings by selling (more) breeding cattle"**
R'=REJECTION VALUE, W'=WEAKENING VALUE (TO CALCULATE MAX. OF INDIVIDUAL SCORE)



ADDITIONAL CONDITIONS AND DEMANDS:.....**none**....................

THE VALUE OF THE AVERAGE SCORE WHERE CONCLUSION IS ABSOLUTELY FALSE:	**-2**
THE VALUE OF THE AVERAGE SCORE WHERE YOU ARE INDIFFERENT:	**0**
THE VALUE OF THE AVERAGE SCORE WHERE THE CONCLUSION IS ABSOLUTELY TRUE:	**+3**

Figure 4.1 Standard form of IMAGINE to be filled in by the expert for the conclusion 'Try to increase earnings by selling (more) breeding cattle'. (See text for explanation)

Terms that the expert uses, like weakening and rejection values, are comparable to those terms that were used in the earlier (unstructured) attempt by LEI-DLO as outlined in section 4.1. The method IMAGINE is a formalisation of the problems we faced earlier.

In figure 4.1, the input data for the required variables to infer the conclusion of our example are written down by the expert:

- **milk quota per hectare** (in kg);
- **milk yield** (average milk yield in kg);
- **deviation purchased feed per cow** (difference between the actual and the adjusted standard value for the amount of purchased feed per cow);
- **deviation cattle credits per cow** (difference between the actual and the adjusted standard value for cattle credits minus animal costs per cow).

Some values of these variables support the conclusion. Such values make the conclusion-supporting area, e.g. values for milk yield greater than 7,500 support the conclusion, while values lower than that origin, value (B), counteract the conclusion. The last values make the conclusion-counteracting area. One or more farm data in the conclusion-counteracting area can be *compensated* by the other values in the supporting area. High values for the second input datum in figure 4.1 support the conclusion (sign=+), and high values for the first, third and fourth input data counteract it (sign=-1).

Normally, the support or counteracting increases constantly with the distance from the origin B (see section 4.3). However, in this example exceptions exist. For two input data the expert indicated the presence of *rejection values*. Values lower than the rejection value 7,125 for milk yield, for example, are allowed but counteract the conclusion increasingly. A farm that realised a low milk yield will have hardly any chance in reaching this conclusion, because the rejection value for this variable makes compensation not very likely. There are also *weakening values*. Such values cause the support (or contribution) from a variable for extreme cases to be limited to that value. The variable would otherwise have too much (positive) influence on the relevance of the conclusion. Very low values for the third and fourth input data, for example, would support the conclusion too strongly. Since the expert did not agree, he has been introducing weakening values for these two variables in figure 4.1.

Another term is the *importance unit*, which expresses the relative importance of each variable. One importance unit is the distance between two rungs in figure 4.1. An example will clarify this. The distance between eight rungs for the first variable is 60,000 - 20,000 = 40,000. One importance unit is then 40,000/8 = 5,000. For the second variable, one importance unit can also easily be calculated: (7,500 - 7,125)/3 = 125. A decrease of 5,000 for the first and an increase for the second variable of 125 are equally important for this conclusion.

The expert indicated that 'additional conditions and demands' are none.

The standard form in figure 4.1 which the expert has filled in, supplies the parameters (variables, importance units origin, etc) for the algorithm of IMAGINE (section 4.3). This algorithm calculates a score for the conclusion based on farm data as input. This score is then transformed to the relevance of the conclusion by means of the three data the expert has expressed on the last lines of the standard form of figure 4.1. At the beginning, these three data are too abstract for the expert to fill in. Default values should be used instead. Only after some testing and experience with that conclusion in the KBS, the transformation can be refined by filling in the actual values.

The time the expert needs to think about one conclusion in this way can take ten to twenty minutes (depending on the difficulty), while the writing down of the necessary information on the standard form will take only a few minutes. Perhaps the most difficult task for the expert is to imagine the relative importance for each variable (importance unit) with reference to the conclusion in question.

The parameterised algorithm and the successive transformation is a model of the expert's knowledge and should be regarded as an approximation of his actual cognitive model. From our experience with IMAGINE, we have reason to believe that this approximation is sufficient for the KBSs we developed.

4.3 The algorithm of IMAGINE

Since the algorithm of IMAGINE is complicated, its theoretic description is illustrated by means of the example from GLOBAL-DETECTOR (Conclusion_8, see section 4.2 and figure 4.1). Especially the third variable of this example, deviation of cattle credits per cow (CC/cow), is used for this purpose.

In general terms, when n possible input data and m possible conclusions are given, the expert tries to find how conclusion i corresponds with a mapping f_i from the values of the input data x_j on the set A of possibilities for relevance, i.e.,

$$f_i(x_1,x_2,\dots,x_j,\dots,x_n) \rightarrow A \qquad (4.1)$$

The relevance of this conclusion is a real number between -100 (absolutely false) and +100 (absolutely true), or stated otherwise, the higher the value of A the more relevant the conclusion is. The expert is indifferent at value zero.

The cognitive model of the expert is represented as f_i.

The data x_j might have a different origin:

- data from an account;
- performance figures which are derived from or composed of data from an account;
- the difference between realised and farm-adjusted standard (FAS) values (see chapter 2). The variable CC/cow is an example for this type;
- the relevance of another conclusion.

IMAGINE approximates f_i, where the approximation should be easy to derive and the conclusion based on the approximations should agree with the expert's conclusion as much as possible. Our approximation of f_i is obtained in two steps. In step 1 we calculate a real value $f^*_i(x_1,x_2,..,x_n)$ (average score) depending on the values $x_1,x_2,..,x_n$ of the input data. The algorithm will be presented after step 2 has been described.

In step 2 this value is mapped into the interval [-100,100] by a function $g_i(\alpha_i,\beta_i,\gamma_i)$, where α_i,β_i and γ_i are parameters supplied by the expert. The function g_i is described in expression (4.2) and illustrated in figure 4.2. The value 2 is used for both a and b in figure 4.2.

$$
\begin{array}{ll}
g(x) = -100 \text{ for } x \le \alpha_i & | \\
g(x) = 100*((x-\alpha_i)/(\beta_i-\alpha_i))^a - 100 \text{ for } \alpha_i < x \le \beta_i & | \\
g(x) = 100 - 100*((\gamma_i-x)/(\gamma_i-\beta_i))^b \text{ for } \beta_i < x \le \gamma_i & | \\
g(x) = 100 \text{ for } x > \gamma_i & |
\end{array}
\qquad (4.2)
$$

As shown in figure 4.2, α_i and γ_i are the values of $f^*_i(x_1,x_2,..,x_n)$ where the conclusion is absolutely false (-100) and true (+100) respectively. At β_i the expert is indifferent concerning the relevance of the conclusion.

<u>EXAMPLE:</u> In the first step, the value for $f^*_8(x_1,x_2,x_3,x_4)$ is calculated. The value of this average score is 0.575, as we shall see in section 4.3.1. This value is meaningless. What we want to know is the relevance of the conclusion. This can be calculated by expression (4.2). For α_8,β_8 and γ_8, the expert has chosen -2, 0 and +3 respectively (expert's knowledge, see figure 4.1). From expression (4.2), it can be calculated that g_8 = 34.7. This makes the conclusion 'slightly relevant'.

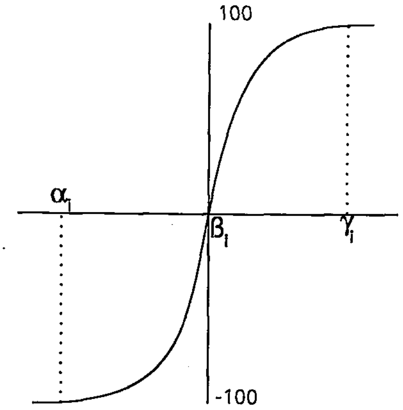


Figure 4.2 Graphical presentation of function $g_i(\alpha_i,\beta_i,\gamma_i)$ from (4.2); a = b = 2

Quadratic terms are used in figure 4.2 (a = b = 2) and in the domain where IMAGINE is applied. An expert can use a value different from 2 when he is not satisfied with the quadratic terms. The curve shown in figure 4.2 will then become more linear (a,b < 2) or more curved (a,b > 2).

The calculation of the average score is performed in step 1. In this step we assume independence of the input data, i.e. the contribution of every input data to the final value $f^*_i(x_1,x_2,..,x_n)$ is determined by the value of the input datum alone, irrespectively of other values. *Compensation*, which is the main characteristic of the method IMAGINE, is possible between the contributions of input data.

To model this, we first scale every input datum j for conclusion i by a scaling factor I_{ij}. Note that this depends on i and j, i=1,..m, j=1,..,n. I_{ij} is called an *importance unit* and is supplied by the expert (see section 4.2). The idea is that one unit of the scaled input datum is as important as one unit of another scaled input datum.

<u>EXAMPLE:</u> The importance unit $I_{8,4}$ for CC/cow can be calculated from figure 4.1. The distance between two rungs is 25. A decrease of 25 is equally important as an increase of 125 for milk yield.

The role of the expert is essential. His task is to provide the parameters required for the functions of the algorithm. Much about his role has been described in section 4.2. His scaling of the importance units (I_{ij}), which express the relative importance of each condition, is knowledge-intensive. They must be scaled with respect to each other.

The value $f^*_i(x_1,x_2,..,x_n)$ is defined by

$$f^*_i(x_1,x_2,..,x_n) = 1/n * \sum_{j=1}^{n} f_{ij}(x_j/I_{ij}) \tag{4.3}$$

where f_{ij} is a function of the scaled input data to the real numbers. The outcome or result of f_{ij} is the individual contribution for input datum j to conclusion i, which will be called *individual score* (with respect to the conclusion at hand) for that datum.

The functions f_{ij}, i=1,..,m, j=1,..,n which calculate individual scores may be application domain dependent, because the method IMAGINE and its functions are especially developed for the KBS GLOBAL-DETECTOR (chapter 6). To apply the method IMAGINE in other domains, the functions f_{ij} might require (minor) modifications. However, the development of other KBSs (prototypes) at LEI-DLO has been done without modifications of functions (e.g. Schakenraad et al., 1994).

For our application domain (GLOBAL-DETECTOR) we have constructed a class of functions which is parameterised by the following parameters that are supplied by the expert (see section 4.2):

- the *sign*; +1 means that the function is monotonously increasing, i.e., higher values support the conclusion;
  -1 means that the function is monotonously decreasing, i.e., smaller values support the conclusion.

EXAMPLE: the sign of the variable CC/cow is -1, since lower values support the conclusion.

- the *origin* B; the value for which the function has value zero. This value is expressed in importance units.

EXAMPLE: the origin for CC/cow is -100. Without any other information, the expert is indifferent about the relevance of Conclusion_8 when CC/cow is -100. For the forthcoming calculation this value is expressed in importance units: $B_{8,4} = -100/I_{8,4} = -4$.

- the *rejection value* R counteracts the conclusion increasingly with increasing (when sign=+1) or decreasing (when sign=-1) values of the variable. The rejection value determines the function f_{ij} for values less than or equal to the origin B for sign=+1 and for values greater than or equal to B in case sign=-1.
  Values for x (a farm value), B and R are expressed in importance units. If no rejection value is given, the function is by definition linear for all x in the counteracting area

  $f_{ij}(x) = (x - B)$ for sign=+1 (4.4)
  $f_{ij}(x) = (B - x)$ for sign=-1 (4.5)

  When no weakening value is given (see below), these functions apply also to all values of x in the supporting area. If a rejection value R is given, and sign=+1, the function is given by

  $f_{ij}(x) = (x - B)$ for $R \leq x \leq B$
  $f_{ij}(x) = (x - B) - ½(x - R)^2 + ½(x - R)^3$ for $x<R$ (4.6)

  Expression (4.6) is illustrated in figure 4.3a.
  If sign=-1, then the rejection value R is greater than the origin B. The function is then given by (4.7) and illustrated in figure 4.3b

  $f_{ij}(x) = (B - x)$ for $B \leq x \leq R$
  $f_{ij}(x) = (B - x) - ½(R - x)^2 + ½(R - x)^3$ for $x>R$ (4.7)

EXAMPLE: The expert has indicated a rejection for CC/cow at position -50, indicating that much higher values are very unfavourable for the relevance of the conclusion. Expressed in importance units: -50/25 = -2. As shown in figure 4.1, high values for CC/cow counteract the conclusion because the sign=-1. When the farm datum for CC/cow is -85, or -85/25 = -3.4 importance units, the contribution to the conclusion is negative: (B - x) = (-4 - -3.4) = -0.6. When we take another farm as an example, where the value is 0 for CC/cow (x = 0), the contribution is very negative: $(B - x) - ½(R - x)^2 + ½(R - x)^3 = (-4 - 0) - ½(-2 - 0)^2 + ½(-2 - 0)^3 = -10$. This low individual score is of great influence on the conclusion and can hardly be compensated by the other variables. For a value of +100 for CC/cow this is even impossible (individual score is -134). This effect is also illustrated in figure 4.5d.

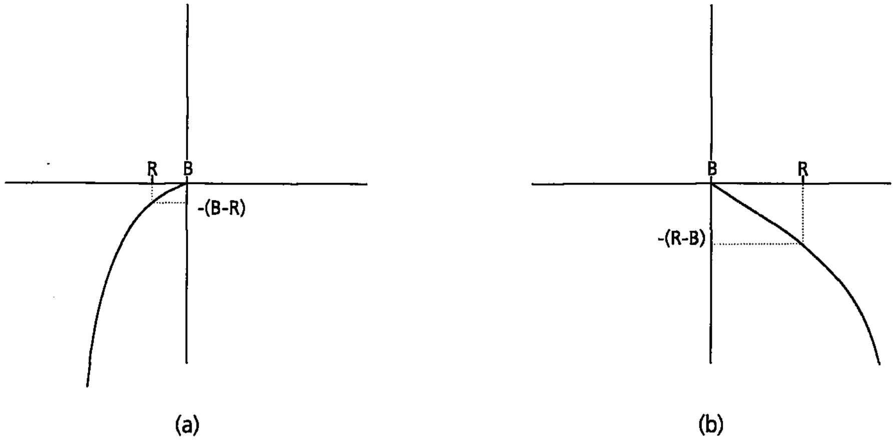


Figure 4.3 Illustration of $f_{ij}(x)$ at the presence of rejection value R, when the sign is +1 (a) and when the sign is -1 (b)

- the *weakening value* W weakens the support for a conclusion when the variable would otherwise have too much influence. The weakening value determines the function f_{ij} for values larger than or equal to the origin B for sign=+1 and for values less than or equal to B in case sign=-1.
  Values for x (a farm value), B and W are expressed in importance units. If no weakening value is given, expressions (4.4) and (4.5) are applied for all x in the supporting area.

  If a weakening value W is given, and sign=+1, the function is given by

  $$f_{ij}(x) = (W - B)*(x - B) / ((W - B)+(x - B)) \quad \text{for } x \geq B \tag{4.8}$$

  Expression (4.8) is illustrated in figure 4.4a.
  If sign=-1, then the weakening value W is smaller than the origin B. The function is then given by (4.9) and illustrated in figure 4.4b

  $$f_{ij}(x) = (B - W)*(B - x) / ((B - W)+(B - x)) \quad \text{for } x \leq B \tag{4.9}$$

  The value of (W - B)*sign is the maximum positive contribution or individual score that can be obtained when $x \rightarrow \infty$ (sign=+1) or $x \rightarrow -\infty$ (sign=-1), as shown in figure 4.4.

EXAMPLE: The expert has indicated a weakening value for CC/cow at position -400. Expressed in importance units: -400/25 = -16. With a farm datum of -229 for CC/cow (x = -229/25 = -9.16), the individual score can be calculated with expression (4.9): (-4 - -16)*(-4 - -9.16) / (-4 - -16)+ (-4 - -9.16) = 3.6. The maximum value that can be obtained is: (W-B)*sign = (-16 - -4)*-1 = 12. This effect is also illustrated in figure 4.5d.

The functions that deal with the weakening value, (4.8) and (4.9), are particular forms of a complementary hyperbolic or Michaelis-Menten function

$$f(z) = A * z / (a + z) \tag{4.10}$$

For IMAGINE, the coefficients in this function are: A = a = (W - B) *sign, and z = (x - B)*sign. The maximum individual score is (W - B)*sign. If x = W, x < W and x > W respectively, then the individual score is exactly half, more than half and less than half respectively as it would be without weakening. In this way, the functions in (4.8) and (4.9) contain fine characteristics which make them understandable and easy to work with.

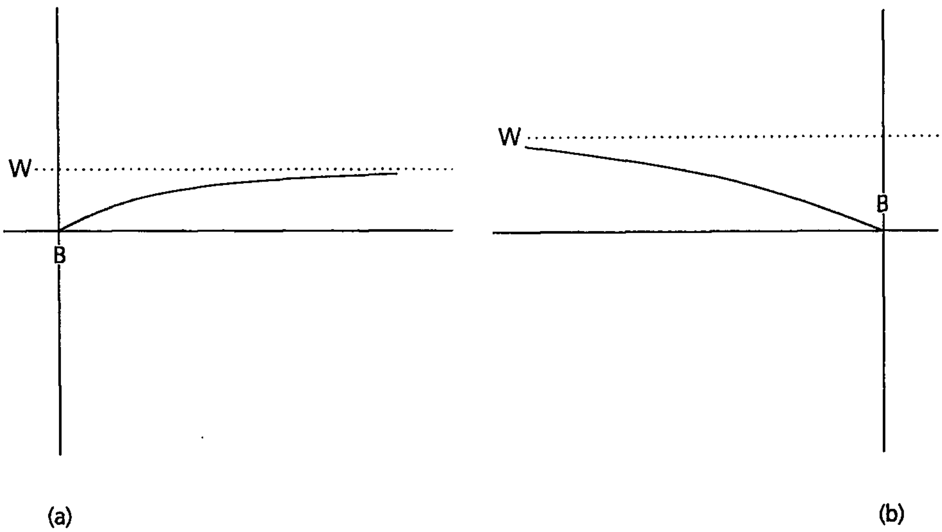


Figure 4.4 Illustration of $f_{ij}(x)$ at the presence of weakening value W, when the sign is +1 (a) and when the sign is -1 (b)

4.3.1 Extended example

The algorithm is embedded in the tool of IMAGINE, developed by LEI-DLO. The contents on the standard forms of the conclusions can be implemented rather fast and easy in the knowledge base of the KBS by the knowledge engineer (or expert). Maintenance of the knowledge by the expert can be performed easily and quickly, he just has to change some data on the standard form (figure 4.1). The time required to implement these modification is negligible.

After the KBS is loaded, farm data from individual farms can be read in the programme followed by the inference (calculation) of the relevance of the conclusions.

For the conclusion in our example, Conclusion_8: **'Try to increase earnings by selling (more) breeding cattle'**, the individual scores for each input datum can be calculated with the algorithm of IMAGINE after the farm data from farm F have been read in the programme. To understand the following calculations, the standard form in figure 4.1 should be consulted.

For the first variable, **milk quota per hectare**, the following information can be obtained from the standard form:

- sign=-1, smaller values of milk quota per hectare support the conclusion;
- the importance unit $I_{8,1}$ = (60,000 - 20,000)/8 = 5,000 kgs;
- when the milk quota per hectare is 20,000, the expert indicated that the value of the score is zero. The value for B is then expressed in importance units to perform the calculations: B = 20,000/$I_{8,1}$ = 4 importance units;
- weakening and rejection values are not indicated by the expert.

The value on farm F for milk quota is 12,584. Also expressed in importance units: F = 12,584/$I_{8,1}$ = 2.5. From (4.5) the individual score can easily be calculated when F is filled in for x: (B - x) = (4 - 2.5) = <u>1.5</u>. The calculation is illustrated in figure 4.5a. The positions of farm value F and origin B are indicated. Notice from this figure that the score decreases constantly by increasing values of the variable, because the sign is -1 and rejection and weakening values are absent. The slope of the line is always sign/$I_{8,1}$ = -0.0002.

The expert indicated a rejection value for **milk yield**, the second variable. Necessary values for the calculation stem from the standard form:

- sign=+1, higher values for milk yield support the conclusion;
- the important unit $I_{8,2}$ = (7,500 - 7,125)/3 = 125 kgs;
- the origin B = 7,500/$I_{8,2}$ = 60 importance units;
- the rejection value R = 7,125/$I_{8,2}$ = 57 importance units. A weakening value is not given.

The value on farm F for milk yield is 7,000, or 7,000/$I_{8,2}$ = 56 importance units.

From expression (4.6), the individual score is: (56 - 60) - ½(56 - 57)^2 + ½(56 - 57)^3 = <u>-5.0</u>. The calculation is illustrated in figure 4.5b, where the positions of F, R and B are shown. Notice that low values for milk yield result in very low individual scores.

For the third variable, **dev. purchased feed per cow**, a weakening value is at issue. Values on the standard form are:

- sign=-1, lower values support the conclusion;
- the important unit $I_{8,3}$ = (0 - -400)/5 = 80;
- the origin B = 0/$I_{8,3}$ = 0;
- the weakening value W = -400/$I_{8,3}$ = -5 importance units. The maximum positive contribution of this variable (=maximum individual score) is also indicated by this weakening value: (W - B)*sign = 5. A rejection value is not given.

The value on farm F for this variable is -316, or -316/$I_{8,3}$ = -3.95 importance units. From (4.9), the individual score is: (0 - -5)*(0 - -3.95) / (0 - -5)+(0 - -3.95) = <u>2.2</u>. The calculation is illustrated in figure 4.5c, where the positions of F, B and W are shown. The position of W marks the maximum individual score to be obtained: 5.

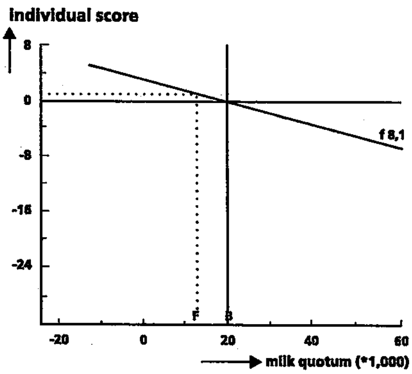


Figure 4.5a Calculation of individual scores for milk quota per hectare (see text for explanation)

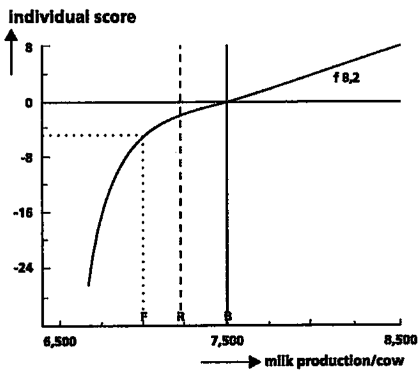


Figure 4.5b Calculation individual scores for milk yield (see text for explanation)

Both rejection and weakening value are indicated by the expert for the fourth variable, **dev. cattle credits per cow**. The values from the standard form are:

- sign=-1, lower values support the conclusion;
- the importance unit $I_{8,4}$ = (-50 - -100)/2 = 25;
- the origin B = -100/$I_{8,4}$ = -4;
- the weakening value W = -400/$I_{8,4}$ = -16. The maximum positive contribution of the variable is (W - B)*sign = 12;
- the rejection value R = -50/$I_{8,4}$ = -2.

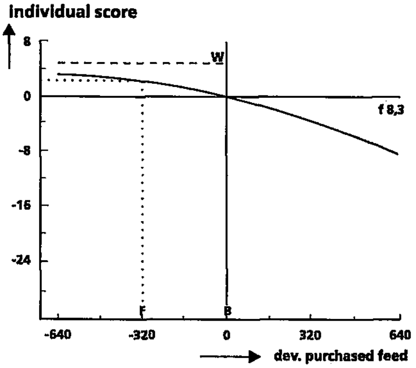


Figure 4.5c Calculation of individual scores for dev. purchased feed (see text for explanation)

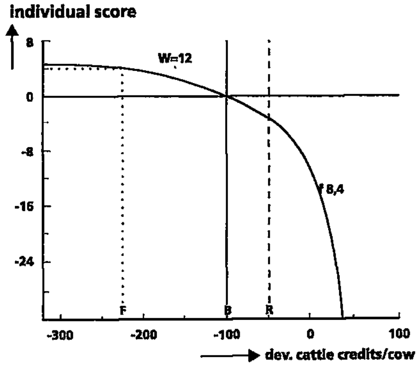


Figure 4.5d Calculation individual scores for dev. cattle credits (see text for explanation)

The farm value F is -229, or $-229/I_{8,4}$ = -9.16 importance units. Because this farm datum is situated in the supporting area (with sign=-1; F<B), the individual score for this variable can be calculated with expression (4.9): (-4 - -16)*(-4 - -9.16) / (-4 - -16)+(-4 - -9.16) = <u>3.6</u>. The calculation is illustrated in figure 4.5d, where the positions of F, B and R are shown. The position of W cannot be shown in this figure, but is just indicated by W=12. Notice the effect of the weakening and rejection value on the individual score. If the farm value is left from B, as in this example, then the weakening value is at issue. Between B and R the individual score is calculated linearly. If the farm value is right from (or higher than) R, the rejection value is at issue.

Table 4.1 summarises the results from the previous calculations of the individual scores from farm F for Conclusion_8.

Table 4.1 Individual scores from farm F for Conclusion_8 (example)

Variable (input datum)	Value from farm F	Individual score
Milk quota per hectare	12,584	+1.5
Milk yield	7,000	-5.0
Dev. purchased feed per cow	-316	+2.2
Dev. cattle credits per cow	-229	+3.6

From (4.3) the final value $f^*_8(x_1,x_2,x_3,x_4)$ is determined by the contribution of every input datum (individual score): (+1.5 + -5.0 + +2.2 + +3.6)/4 = 0.575.

In this way, each individual score is treated independently and compensation is possible. The final value will be called *average score*, since it is the average of the individual scores.

So far, step 1 is illustrated. The average score of 0.575 is in a sense meaningless. The value has to be converted to the relevance of the conclusion, which is meaningful. This is done in step 2 by the function $g_i(\alpha_i,\beta_i,\gamma_i)$. Step 2 has been explained at the beginning of section 4.3. The expert has assigned the necessary values for the calculation: α_i=-2, β_i=0, γ_i=3 (see figure 4.1, last lines).

With (4.2) the relevance of Conclusion_8 for farm F can be calculated: 100 - 100*((3-0.575)/(3-0))^2 = 34.7. On a scale of -100 (very irrelevant) to +100 (very relevant), this relevance can be interpreted as: Conclusion_8 is 'slightly relevant'. In this example, the negative support of milk yield for Conclusion_8 is more than fully compensated by the three other factors. When the value of the average score would have been be zero, the expert is indifferent when the conclusion is true or not (relevance = 0). When the average score was 2.6, -0.7 and -2.5 respectively, the relevance would be 98 (very relevant), -78 (rather irrelevant) and -100 (very irrelevant) respectively.

Figure 4.6 illustrates the calculation of the relevance from the average score.

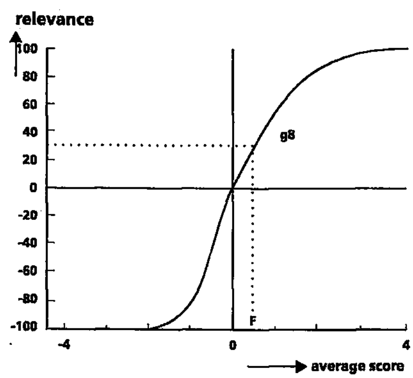


Figure 4.6 Calculation of the relevance of the conclusion (see text for explanation)

Variables that are strict conditional or crisp may also take part in IMAGINE. The method is not restricted to merely 'fuzzy' variables, like the ones used in our example.

4.4 The presentation of the conclusions

In the KBS GLOBAL-DETECTOR, the relevance of all conclusions are calculated with the same method and algorithm. This means that (4.4) to (4.9) are used for the calculation of the individual scores, and that (4.2) and (4.3) are used to calculate the relevance from the average of the individual scores. All conclusions are presented to the user in sorted order, making discrimination possible.

The user is allowed to ask the system how a certain conclusion has been reached. Firstly, an easy readable text is shown to him. This text, which was formulated by the expert during the development, informs the user how a certain conclusion generally is inferred.

Secondly, very detailed and quantitative information is presented to the user (figure 4.7). Again, the same example from section 4.2 is taken.

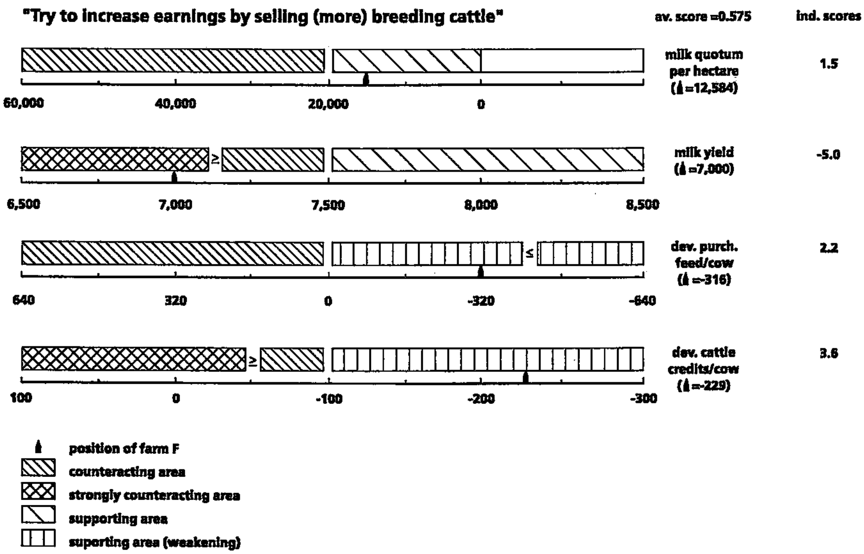


Figure 4.7 Detailed explanation facility concerning Conclusion_8 from farm F

On top of figure 4.7, the name and the average score of the conclusion is presented. The four variables, which are part of the conclusion, are shown on the right side, with their values from farm F. The influence of the variable on the conclusion is expressed by the bar. The values in the centre (20,000, 7,500, etc) are the values that determine the origin B, or the values where the expert is indifferent. The right side of the bar is the supporting area, the other side is the area that counteracts the conclusion. The farmer's value is also shown. The milk quota per hectare on farm F (12,584) is supporting the conclusion, resulting in an individual score of 1.5 (shown at the very right side). See Section 4.3.1 and figure 4.5a for the calculation of this score.

The rejection value is indicated by ≥. The symbol ≤ is the position of the weakening value W. At this position the individual score is half of the score without weakening (see expression (4.10)). For additional explanation of figure 4.7, the reader is referred to Section 4.3.

The information the expert has written down on the standard form (figure 4.1), appears almost identically on the screen (figure 4.7). With the easy readable text shown to him, the user gains an insight into the cognitive model of the expert regarding the conclusion.

4.5 Explanation facilities: IMAGINE versus rule-based systems

KBSs based on the method IMAGINE are comparable with traditional rule-based systems, the most used type of KBS. However, a striking difference is the fact that IMAGINE can be used in domains characterised by quantitative and continuous data, that it can handle the problem of combinatorial explosion, and that it can take into account the importance of individual conditions. Another difference is formed by the explanation facilities.

The user experiences the differences with traditional systems most with respect to the explanation facilities. When IMAGINE is used, *all* information concerning a conclusion or solution is gathered on one screen (figure 4.7), very much comparable with the knowledge of the expert on the standard form (figure 4.1). On the condition that the user is fully informed about the method IMAGINE, the information is comprehensive.

Most explanation facilities for traditional rule-based systems are questionable, especially with the existence of many related rules. Even for the knowledge engineer and/or the expert, the presented rule or a chain of rules is not always clear, due to several reasons:

1. the choice of the conditions as well as (borderline) values are not explained;
2. the particular rule is part of a structure, which can, if possible, only be revealed by extensive use of the explanation facilities. Why the structure is set up in that way, may not be explained either;
3. the context in which the presented rule did fire may not be known;
4. necessary arithmetic calculations resulting in a value for a condition are mostly not shown.

Remarks on the explanation facilities are also reported by Jackson (1986). When the user does not get enough information concerning the rationale behind the drawn conclusions, it may be assumed that the acceptance of conclusions will not be high.

For the user the results must be easy to interpret and reliable at the same time. A drawback of IMAGINE is that the user has to be informed to understand the explanation facilities of figure 4.7. Some farmers of the test group could interpret the results very well with the built-in help facilities from the computer programme as only support. Others had trouble understanding them, which could lead to non-acceptance of the system. The systems built with IMAGINE are not only to be used by farmers, but especially by advisors who can explain the results from IMAGINE. The advisors can additionally discuss the results with farmers.

This major drawback of IMAGINE is absent in the method FUZZY-DETECTOR, to be presented in the next chapter. This new method is grounded on the ideas of IMAGINE and the fuzzy set theory. FUZZY-DETECTOR is less detailed and presumably not as reliable as IMAGINE, but its explanation of the outcome is quite understandable with little or no support.

4.6 Discussion: prospects and limitations of IMAGINE

All functions described in section 4.3, which are in fact the roots of IMAGINE, are developed with the expert for the domain at hand. Their characteristics are chosen for pragmatic reasons. The method is a trade-off between ease of use and reliability. It must be understandable and workable for both expert and user. If one would ask the expert to give (un)certainties, dependencies between concepts, etc, the problem would be too complex for the latter to oversee. This might even go beyond the cognitive capabilities of the expert's mind. For the end user, the presentation shown must be understandable and not too theoretical.

Although the application of the functions proved successful during the development of some systems, it must be stressed that these functions are not rigid ones. Knowledge engineers working with IMAGINE in other domains are free to adapt these or use others. No adaptions are made by the knowledge engineer when the method was applied for the identification of the style of farming (chapter 8).

The process of 'backward knowledge acquisition' should be regarded as a model driven approach (section 4.2). The expert concentrates on a conclusion and uses the method and standard form of IMAGINE (figure 4.1) as a model to put down his knowledge about the concepts or data in relation to the conclusion in a structured way. Such an acquisition of knowledge is preferable to a data driven approach (Breuker and Wielinga, 1989).

The model-based knowledge acquisition by IMAGINE appears to be artificial. The nature of the method with its algorithm and standard form is in fact normatively thrusted upon the expert, who must be fully informed in great detail before start. The proposed framework or model does not make his job an easy one, especially the judgement of the importance of each individual variable with regard to the other ones. It is definitively plausible that this is not the way he thinks or solves problems. The result is merely a simplified cognitive model of the expert.

But notwithstanding all that, the expert's experience with the application of IMAGINE at LEI-DLO is positive. From our experience, we have reason to believe that the introspective process of 'backward knowledge acquisition', as well as the resulting representations, model the way of thinking in this domain satisfactorily enough.

It must be emphasised that the user gets the knowledge presented in the same manner the expert had in mind when he filled in the standard form. In fact, there is a 'direct' step. The knowledge engineer has no (negative) influence on this. So, the role of the knowledge engineer is very restricted. His main task is a clear explanation of the method IMAGINE and how it should be used. After this is done with great care, his presence during the process of knowledge acquisition is actually not necessary. In this context, Hayes-Roth et al. (1983) note: 'The knowledge engineer's job is to act as a go-between to help build an expert system. Since the knowledge engineer has far less knowledge of the domain than the expert, however, communication problems impede the process of transferring expertise into a programme'.

4.7 Conclusion

After experiences with the development of a few KBSs, it can be asserted that IMAGINE is a method for the model-based acquisition, representation, presentation and maintenance of knowledge in a fast, effective and straightforward way. Especially in a domain dominated by quantitative and continuous variables. Referring to our objective, IMAGINE meets the problem of combinatorial explosion quite satisfactorily because of the introduction of smooth or fuzzy boundaries and the allowance of compensation between different concepts.

The arithmetic functions of IMAGINE are essential for the method. The relevance or truth content for each conclusion can be calculated with these functions. All conclusions are presented to the user in sorted order with respect to relevance, making discrimination possible.

The method and especially the meaning of the arithmetic functions must be explained thoroughly to the expert before knowledge acquisition can start. After this has been done, the time the knowledge engineer has to spend is minimal since the expert can do his task independently.

5. FUZZY-DETECTOR: FUZZY SETS FOR PERFORMANCE EVALUATION UNDER UNCERTAINTY

'... the world is fuzzy, therefore our mathematics should also be fuzzy.' (French, 1984)

5.1 Introduction

In chapter 4, IMAGINE was presented as a method to build knowledge-based systems (KBSs) for domains where the majority of variables is continuous. In such domains an unmanageable number of situations exist when these continuous variables are made discrete. This problem of combinatorial explosion was met by the introduction of smooth or fuzzy boundaries and the possibility of compensation between different concepts.

But users who are less well-informed about IMAGINE, have trouble understanding the explanation facilities of systems developed with this method. These users urged on the developers the necessity of a less quantitative approach and as a result more clearness.

Limitation of IMAGINE is also the disability to deal with uncertain and qualitative data. To take into account farmer's goals, wishes and styles of farming and to extend to environmental problems, data are often incomplete, uncertain and difficult to handle, and in many cases information on probabilities is lacking (Janssen, 1991). It is to be expected that such data become increasingly important in (knowledge-based) computer programmes.

The objective of this chapter is to describe the method FUZZY-DETECTOR, which tackles the problem of dealing with qualitative and uncertain data. And what is more, a clearer explanation facility for the user is an advantageous side-effect.

The forthcoming presentation of this new method is illustrated with an example right from the start. The text will be about the aspect of uncertainty in the knowledge and the data, a short introduction of the fuzzy set theory, the method FUZZY-DETECTOR in detail and the theory which lies at the root of the method at issue.

It must be stressed that IMAGINE by no means is inferior to FUZZY-DETECTOR. Which method to apply depends mainly on the characteristics of the domain (quantitative versus qualitative) and the required understandability.

5.2 General outline

In this section we will provide a general outline of the method FUZZY-DETECTOR. In the sections below we will describe the details. The central issue in FUZZY-DETECTOR is how to handle uncertainty. The uncertainty we refer to is not uncertainty in the probabilistic sense but uncertainty with respect to classifying an element as belonging to a set due to the vague and imprecise definition of the set, i.e., we refer to uncertainty in the sense of fuzzy set theory 1). Since most readers will not be familiar with the fuzzy set theory we will explain this through an example.

5.2.1 A short introduction to the fuzzy set theory

Let us consider the set of very large persons. An ordinary definition could define this set as the set of all persons larger than or equal to 1.95 metres. Being an element of the set is a yes or no question; or putting it differently, the membership function (MSF) which assigns to each element a value can have two possible values, namely 0 indicating that the element does not belong to the set, and 1 indicating that the element does belong to the set. A person with a height of 1.94 metres will have a MSF value 0, and a person with a height of 1.96 metres will have a MSF value of 1. Most people will agree that such a big gap in MSF value for the two persons is a bit strange since there is hardly any difference in height.

The problem is that we have tried to make precise such a vague concept as the largeness in height. This can be avoided if we would allow the MSF to take any value between 0 and 1. A person with a height of 1.85 metres would have a MSF value of say 0.85 indicating that it is almost a very large person. This is exactly the way fuzzy set theory handles sets defined in linguistic terms.

Formally a fuzzy set S is defined by an ordinary set X, called the ground set, and a MSF $\mu_S : X \to [0,1]$ which assigns to each element in X a value in [0,1].

A possible MSF for the set of very large persons is given in figure 5.1 below. In the figure on the left, the ground set is [1.00,2.00]. In FUZZY-DETECTOR, we assume that the ground set of all fuzzy sets is [0,1]. This is without loss of generality since the ground set of every fuzzy set relevant for our application domain is bounded, and hence can

---

1) The fuzzy set theory was first introduced by Zadeh (1965). Since then, a vast amount of articles and applications on this area have emerged. Comprehensive literature exist on basic theoretical aspects of fuzzy sets, especially Zimmermann (1991) is both extensive and accessible.

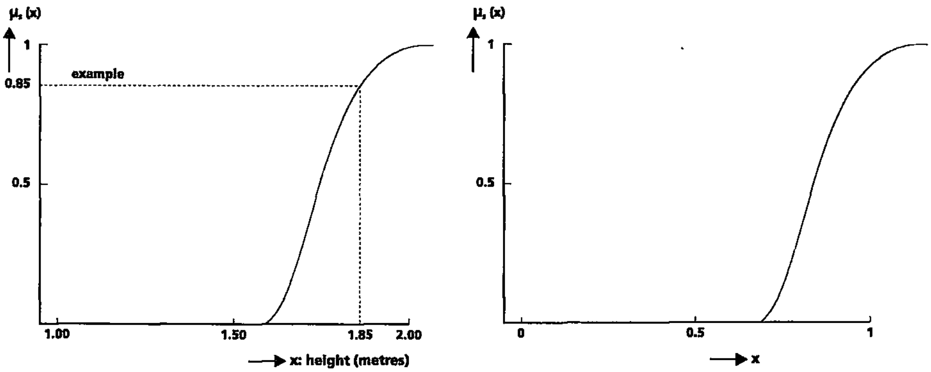


Figure 5.1 MSFs for the set 'very large persons' with different ground sets

be mapped into [0,1] by linear transformation. In figure 5.1, the interval [1.00,2.00] of the left figure has been transformed 1) linearly to [0,1] in the right figure.

For this example the linguistic term 'very large' has been used. Many different terms are used in FUZZY-DETECTOR. The MSFs which go along with these terms are difficult to understand. Suppose we have a certain concept, which value might be one out of the set ('very bad', 'bad', 'rather bad', 'below average', 'average', 'above average', 'fair', 'good', 'very good'). Such concept can be described with a MSF. Figure 5.2 shows the MSFs when the linguistic value of the concept is either 'fair', 'good' or 'very good'.

A simple analogy will be used to explain figure 5.2. Suppose X is the set of possible report marks used at schools in a fictitious country. The marks $x \varepsilon X$ the students receive range from 0 ('very bad') to 1 ('very good'). When the mark of a particular student is 0.8, one might call this 'good', while a mark of 0.9 might be denoted as 'good' or 'very good'. As can be seen from figure 5.2, for the rating (or mark in our analogy) 0.8, its member in the set 'good' is 1 and its member in the set 'very

---

1) To apply FUZZY-DETECTOR, the ground set X is the result of the transformation from the original values. Each value in X is calculated from original values of x by

$$X = \begin{cases} 0 & ; x \leq a \\ (x - a)/(b - a) & ; a < x \leq b \\ 1 & ; x > b \end{cases}$$

The values a and b are the lowest and highest values of interest from the original interval. Values lower than a or greater than b have the same meaning than a or b. If the original set for the length of people expressed in metres has values in [a,b]=[1.00,2.00], then a person with a length in X of 1.75 metres corresponds to a value of (1.75-1.00)/(2.00-1.00)=0.75.

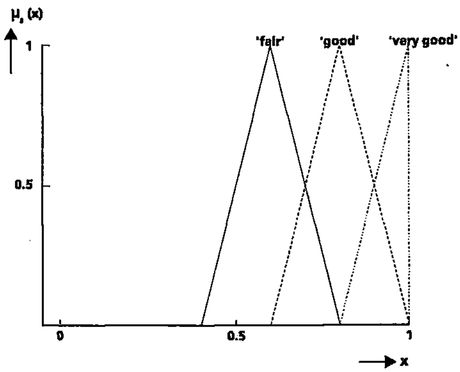


Figure 5.2 MSFs for linguistic terms 'fair', 'good' and 'very good'

good' is 0. A higher rating will decrease the grade of membership in the set 'good', while increasing the grade in the other set. A rating or mark of 0.9 has an equal grade in both sets: 0.5.

MSFs do not have generally accepted shapes. The shapes are different throughout the literature. The functions presented in this chapter are chosen because they fit the application. However, this still has to be validated. Each person who wants to apply the method FUZZY-DETECTOR may use his own shapes of MSFs.

5.2.2 The knowledge base of FUZZY-DETECTOR

Knowledge in FUZZY-DETECTOR takes the form of IF-THEN rules. In the IF part conditions and their importance are stated; in the THEN part the conclusion is stated. Fuzzy sets arise in FUZZY-DETECTOR because farm data and conditions may be expressed in linguistic terms. An example of an IF-THEN rule, that will be used from now on, might clarify this.

EXAMPLE :

IF

(Application low emission technique is <SLIGHT) [VERY_IMPORTANT]

AND

(Time of slurry application is <BAD) [IMPORTANT]

AND

(Storage capacity of manure is <LOW) [MODERATELY_IMPORTANT]

THEN

Bad emission-conscious management

The conclusion 'Bad emission-conscious management' 1) in the THEN-part has a value of relevance for a particular farm calculated from the degree to which the farm data satisfy the conditions. These conditions have a different degree of importance (e.g. VERY_IMPORTANT for the first condition). This IF-THEN rule from our example can be acquired from a verbal report of an expert:

'On a particular farm, that is situated on a sandy soil, there is talk of bad emission-conscious management when there is only slight or very slight usage of low-emission technique, when the farmer applies the slurry late in the year, and when the storage capacity of manure on the farm is low. Especially the application of low-emission technique is very important for the relevance of this conclusion, it is the most important way to reduce ammonia volatilisation on a farm. The time of application is only a bit less important, while the importance of storage capacity is moderate.'

This IF-THEN rule may be part of the knowledge base for the tool where the method FUZZY-DETECTOR is implemented and will be used to find out to what extent the conclusion 'Bad emission-conscious management' is true for a farm F. For this conclusion the first variable (Application low emission technique) has the condition 'at most slight' (i.e. <SLIGHT) and is very important. When the value for this variable is 'rather slight' on farm F, we have to find to what extent this value matches the condition <SLIGHT. The result of this match together with the accompanying importance (i.e. 'very important') determines the contribution for the conclusion. All steps, from data and conditions to the relevance of the conclusion, are described below.

5.2.3 The use of membership functions (MSFs) in FUZZY-DETECTOR

Fuzzy sets are described by their MSFs. Such functions are used throughout the method FUZZY-DETECTOR. They will be presented according to the different parts or different roles they play in the method.

DATA OF FARM F

Farm data are expressed as MSFs. For our example, we assume the following farm data for farm F:

Application low emission technique is RATHER_SLIGHT [UNCERTAIN]
Time of slurry application is BAD [CERTAIN]
Storage capacity of manure is FAIR [CERTAIN]

---

1) The example is supplied by ing. H.H. Luesink from LEI-DLO. The first condition is restricted to non-sandy soils (Emission = ammonia losses due to volatilisation). The prefixes > and < for the conditions mean 'at least' and 'at most', respectively. So, *<SLIGHT* means 'at most slight' and *<BAD* means 'at most bad'.

The certainty status of farm data indicates how certain the information supplier is about the correctness of the farm data. This is reflected in the MSF of the farm data. In figure 5.3 we give the MSF of the farm datum 'Storage capacity of manure is FAIR' for every possible certainty status. There are two observations to make. First of all, the ground set of all fuzzy sets is [0,1] as we have assumed earlier for FUZZY-DETECTOR. The second observation relates to the shape of MSFs. All MSFs have a trapezium shape and can be completely characterised by the parameters a, b, c, and d, where a<=b<=c<=d. In figure 5.3 (left figure), the values for a, b, c and d are 0.5, 0.6, 0.6 and 0.7 respectively. Note that here a degenerated case occurs because two or more of these parameters are equal (b=c).

The midpoint shall be defined as the centre of the interval of the set for which the membership values equal one (i.e. μ=1). If, for example, the values for b and c are 0.5 and 0.7 respectively, the midpoint is {(0.5+0.7)/2} = 0.6 (see e.g. figure 5.3, right figure). In FUZZY-DETECTOR, the certainty status does not affect the midpoint between b and c; this midpoint does not change when the uncertainty increases. An increase in uncertainty results in an increase of the interval [b,c]. The difference between a and b, and the difference between c and d are affected also (see figure 5.3).

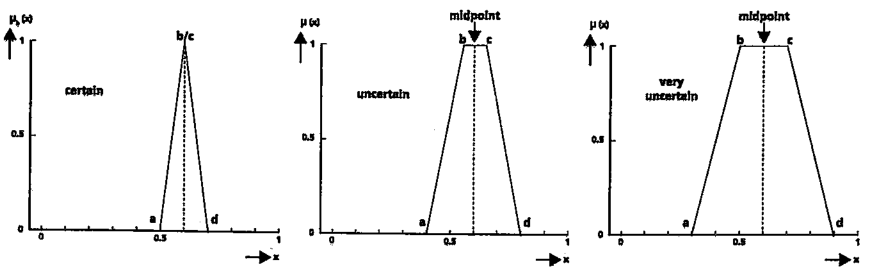


Figure 5.3 Effect of the certainty status on the shape of the MSF for the concept 'Storage capacity of manure is FAIR'

So, with MSFs we can express uncertainty in the farm data. But most data from an account are numeric and not linguistic and without any uncertainty. To apply FUZZY-DETECTOR for the interpretation of farm accounts, such data have to be modelled also. A milk yield per cow may have a value of 7,400. When all values in the interval [5,000, 9,000] are transformed to [0,1], this milk yield would get a rating of 0.6 (calculated as {(7,400-5,000)/(9,000-5,000)}). Figure 5.4 shows the function of this milk yield as farm datum. In this special case there is no uncertainty, the degree of membership is 1 for the rating of 0.6 and 0 for all other ratings. It is a special kind of a MSF for a fuzzy set; the (transformed) milk yield of 0.6 will be called a crisp value.

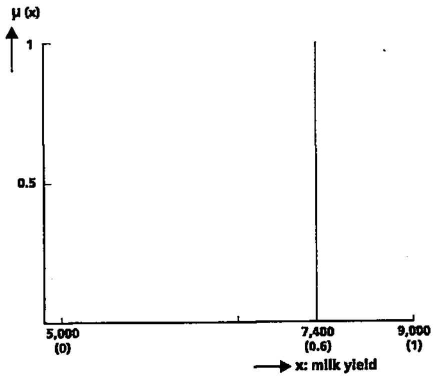


Figure 5.4 A special kind of MSF for 'milk yield per cow = 7,400 (x=0.6)'

CONDITIONS FOR THE CONCLUSION(S)

Figure 5.5 presents S-shaped MSFs for some conditions. The conditions in our example <BAD, <SLIGHT and <LOW, have equal shaped MSFs. <BAD, for example, stands for 'at most bad'. MSFs for 'very bad', 'at most rather bad', 'below average', 'above average', 'at least rather good', 'at least good' and 'very good' are presented as well. MSFs for other linguistic expressions can be inferred from these. For example, the MSF for 'at most bad' in figure 5.5 is equal to the MSFs for 'at most slight' and 'at most low'.

All functions in figure 5.5 can be modelled with the parameterised functions in (5.2), where $0 \le \alpha \le \beta \le \gamma \le 1$. These functions are derived from expression (4.2) in chapter 4. For continuously decreasing and for continuously increasing MSFs respectively, a different set of (related) functions is used in (5.2). At $x=\beta$, the value of $\mu(x)$ is 0.5.

	continuously decreasing:		continuously increasing:		(5.2)
$\mu(x) =$	1	or	0	for $x \le \alpha$	
$\mu(x) =$	$1 - ½*((x-\alpha)/(\beta-\alpha))^2$	or	$½*((x-\alpha)/(\beta-\alpha))^2$	for $\alpha < x \le \beta$	
$\mu(x) =$	$½*((\gamma-x)/(\gamma-\beta))^2$	or	$1 - ½*((\gamma-x)/(\gamma-\beta))^2$	for $\beta < x \le \gamma$	
$\mu(x) =$	0	or	1	for $x > \gamma$	

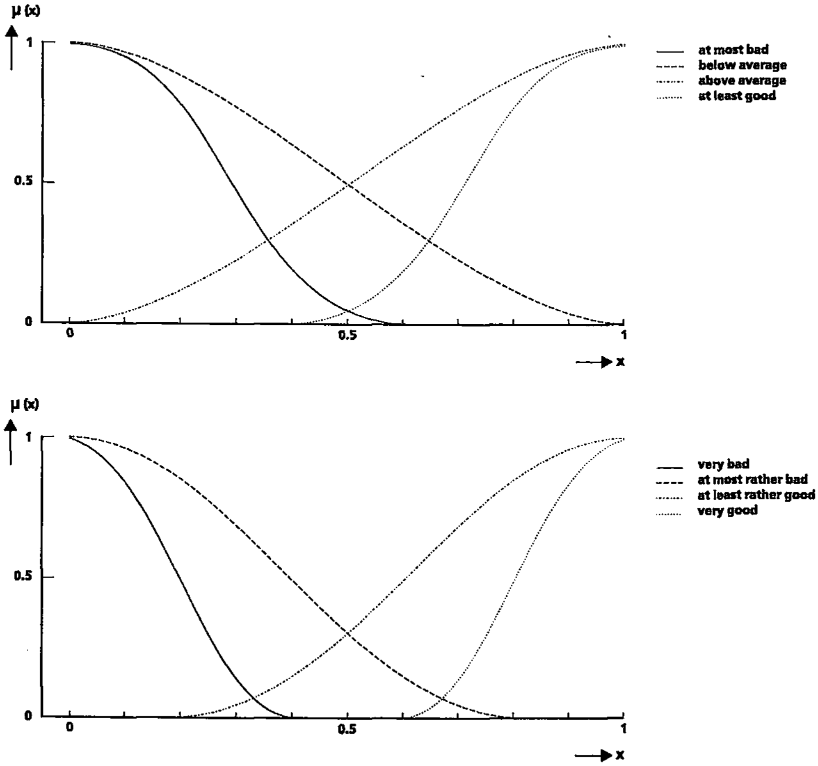


Figure 5.5 S-shaped MSFs used for some conditions

MATCH BETWEEN A FARM DATUM AND A CONDITION

It has to be found out to what extent farm data match the conditions. Given the MSFs of a condition and a corresponding farm datum, we present in section 5.3 a procedure to calculate the MSF of the match between them. The MSF values indicate how well the farm datum satisfies the condition. The better the match, the higher the value of the matching function will be.

This is illustrated in figure 5.6. The MSF of the match between farm datum 'bad' and condition <BAD is shown here. In section 5.3 and figure 5.11 the calculation and the appearance of this MSF is explained.

As can be seen from figure 5.6, the range of the matching function is the ground set [0,1] of a fuzzy set. This range goes from a perfect mismatch (m=0) to a perfect match (m=1). The MSF in this figure is not a perfect match, here at m=0.88 the maximum MSF value is reached (μ(0.88)=1). Due to uncertainty, values around m=0.88 also have high MSF values. It is not always easy to give a description in linguistic terms of this (matching) fuzzy set. The MSF of figure 5.6 could be described as a good match. However, a linguistic description is not necessary for the successive calculations and for the presentation to the user.

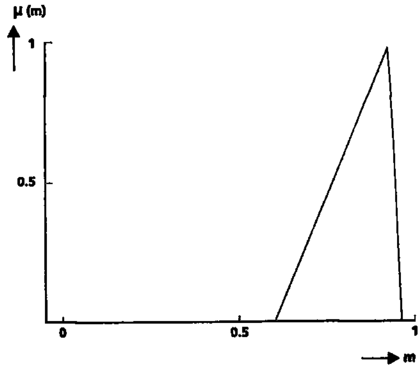


Figure 5.6 MSF for the match between condition <BAD and farm datum 'bad'

IMPORTANCE OF INDIVIDUAL CONDITIONS

We define the relevance function r of a conclusion, e.g. 'Bad emission-conscious management', to be the weighted average of the matching values of all conditions. Here the weights reflect the importance of the conditions, and these are also expert knowledge (see the example of the IF-THEN rule in section 5.2.2). By taking the weighted average, we allow for compensation between matching values of individual conditions by taking into account the relative importance of matchings. A good match for an important condition gives a lot of support to the conclusion. An unimportant condition has little impact on the support of the conclusion.

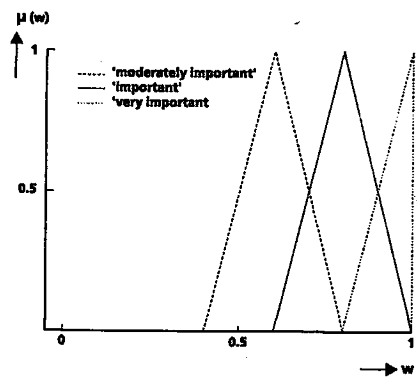


Figure 5.7 MSFs of weights corresponding to importance classes

The weights are elements of a fuzzy set. MSFs are given in figure 5.7 for three importance classes: 'moderately important', 'important' and 'very important'.

RELEVANCE OF THE CONCLUSION (Baas and Kwakernaak, 1977; Kwakernaak, 1979)
The MSFs for the match between conditions and farm data are combined with the MSFs for the corresponding weights according to the method of Baas and Kwakernaak (1977) and the algorithm of Kwakernaak (1979). The objective of these calculations, which will be described in section 5.4, is to find the relevance of a conclusion. This relevance is also a MSF. Figure 5.8 shows the MSF of the relevance of our example (section 5.2.2) for the data of farm F.

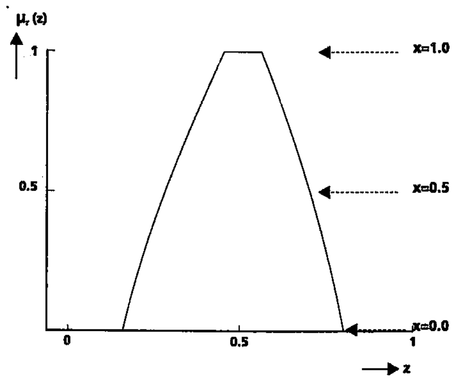


Figure 5.8 MSF of the relevance of the conclusion 'Bad emission-conscious management' for farm F (example)

The range [0,1] of the relevance function r of the conclusion goes from 'very irrelevant' (z=0) to 'very relevant' (z=1). Intermediate values may be 'irrelevant', 'rather irrelevant', 'slightly irrelevant', 'slightly relevant', 'rather relevant', and 'relevant'. The range [0,1] of the relevance function r is the ground set of a fuzzy set. Like the MSF of the matching (figure 5.6), it is not always easy to characterise the corresponding fuzzy set in linguistic terms. To describe the fuzzy set corresponding to the range of the relevance function in linguistic terms, one might concentrate on the values for which the MSF value μ is one, and characterise the set by the corresponding degree of relevance. The relevance in figure 5.8 might then be called something between 'slightly irrelevant' and 'slightly relevant'.

In the output of the computer programme where the method FUZZY-DETECTOR is implemented as a tool, the MSF of r is not completely calculated and presented like in figure 5.8, but only two so-called α-levels are presented like in figure 5.9. The interval of all r values where $\mu(r) \geq \alpha$, is called the level set $L(\alpha)$. The level set L(1) where the value of μ is equal to 1 ($\alpha=1$), has interval [0.44,0.56] (see figure 5.8). The level set L(0.5) where the value of μ is greater than or equal to 0.5 ($\alpha=0.5$), has interval [0.30,0.70]. The level set L(0) where the value of μ is greater than or equal to 0 ($\alpha=0$), always has the ground set [0,1] as interval. For the value of μ greater than 0 the interval is [0.19,0.81].

For the calculation of the MSF for the relevance with the algorithm of Kwakernaak (1979), as to be explained in section 5.4, only the intervals at α-levels of the matching MSF are required. KBSs built with FUZZY-DETECTOR only calculate the intervals at α-levels 1 and 0.5. In this chapter the level where $\mu(r)>0$ is illustrated occasionally.

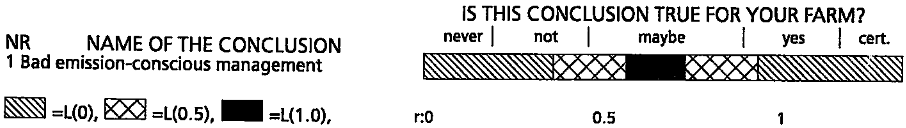


Figure 5.9 Explanation facility for the conclusion 'Bad emission-conscious management' for farm F, by method FUZZY-DETECTOR (cert.= certain). Output from FUZZY-DETECTOR

5.3 Matching condition and farm datum

The MSF μ_C for a condition C describes the extent to which an element $x \varepsilon [0,1]$ satisfies the condition, i.e. the larger $\mu_C(x)$ the better x satisfies the condition. So there is a good match between data element x and condition C whenever $\mu_C(x)$ is close to 1. We call $z=\mu_C(x)$ the matching value of x and say that x supports the matching value z. Of course there may be more than one data element supporting a matching value. We define the matching MSF value $\mu_M(z)$ of a matching value z as the data membership value of the best data element supporting z, i.e. the data element with the largest MSF value of the condition.

The MSF of a (farm) datum is indicated by μ_D (see e.g. figure 5.3), and the MSF of a condition by μ_C (see figure 5.5). A data element x in [0,1] has matching value $\mu_C(x)$ with the MSF of the condition. The MSF μ_M of the match is now defined by

$$\mu_M(z) = \underset{x:\mu_C(x)=z}{\text{supremum}} \{\mu_D(x)\} \qquad (5.2)$$

We illustrate the calculation of the α-levels by our example. The MSF for the first condition of our example in section 5.2.2, 'application of low emission technique is at most slight', is drawn on top of figure 5.10. This condition should be matched with the datum from farm F: 'rather slight' [uncertain]. This MSF is shown in the middle of figure 5.10. The calculation (i.e. matching) may be done at several α-levels. For the explanation of the procedure we will restrict ourselves to α-levels 1 and 0.5 and the level where $\mu>0$. In section 5.2 this restriction has been explained.

From the MSF of the farm datum we determine the interval [x1,x2] with the property that $\mu_D(x) \geq \alpha \Leftrightarrow x \varepsilon [x1,x2]$. For $\alpha=1$ we have for x1 and x2 the values 0.25 and 0.35 respectively, since $\mu_D(0.25)=\mu_D(0.35)=1$. The interval is [0.25,0.35] for $\alpha=1$. The next step is to find the membership values of the condition at this interval. As shown in the topmost graph of figure 5.10, the lowest value is at x2: $\mu_C(x2)=\mu_C(0.35)=0.35=z1$. The highest value is at x1: $\mu_C(x1)=\mu_C(0.25)=0.65=z2$. The membership values $\mu_M(z)$ of the resulting MSF of the match (figure 5.10, bottom) are 1 (i.e. α) at the interval [z1,z2]=[0.35,0.65].

Notice that μ_C is continuously decreasing and that no $x \varepsilon [x1,x2]$ exists where $\mu(x)>\mu(x1)$ or where $\mu(x)<\mu(x2)$. Such a problem is absent in FUZZY-DETECTOR since all MSF of the condition are either continuously decreasing (e.g. 'at the most bad') or continuously increasing (e.g. 'good').

The same procedure has to be followed for other α-levels. For α-level 0.5 the interval [x1,x2] of the farm datum is [0.175,0.425]. From the MSF of the condition it can be derived that the interval for the matching MSF is [z1,z2]=[0.17,0.83] at α-level 0.5. Here $\mu_M(z) \geq 0.5$. Finally, for α-level 0 (or actually a little bit higher than 0) the interval [x1,x2] is (0.1,0.5). This results in interval [z1,z2]=(0.06,0.94) of the matching MSF where $\mu_M(z)>0$. Notice that we make an exception for L(0), because $\mu_M(z) \geq 0$ would result in the ground set [0,1].

To construct the match in figure 5.10, the calculation at a large number of α-levels is required. The match can be described as a match somewhere between 'rather bad' and 'rather good'. The uncertainty makes an exact description not possible.

Let us apply the matching algorithm to the two other conditions and farm data given in section 5.2.2. The figures 5.11 and 5.12 give the MSFs μ_C, μ_D, and the resulting MSF μ_M for the second and third condition of the example. The same procedure can be obtained to get the intervals at various α-levels.

At x1=x2=0.15 the MSF value of the farm datum $\mu_D(x)$ is 1 in figure 5.11. Notice that in this case there is no interval. At x=0.15 the MSF value for the condition $\mu_C(x)$ is 0.88. So z=0.88. The membership value of the matching MSF $\mu_M(z)$ is 1 for z=0.88. For α-level 0.5 the interval [x1,x2] is [0.1,0.2], resulting in an interval for the match of [0.78,0.94]. Figure 5.11 shows the procedure at this α-level. For α-level 0 the interval [x1,x2] is

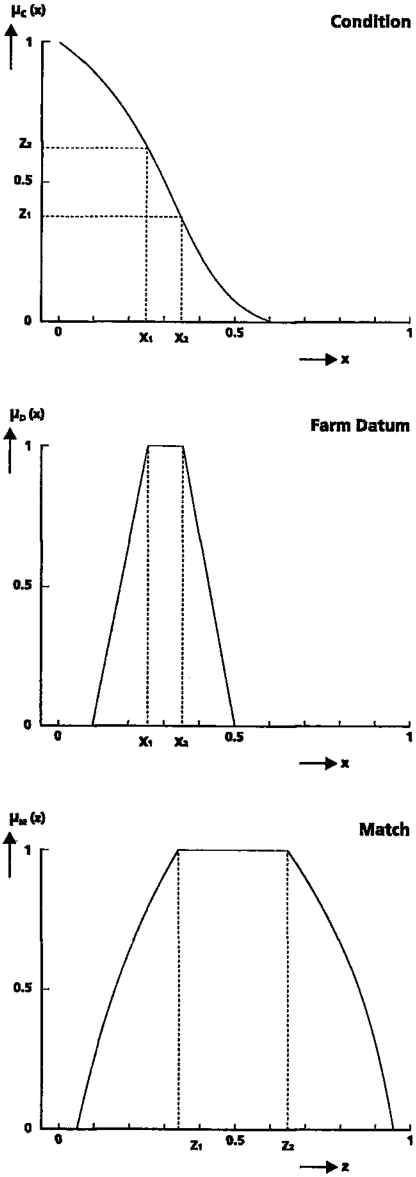


Figure 5.10 *MSFs for condition 'application low emission technique is at most slight', for farm datum 'rather slight', and for the resulting match. See text for explanation*

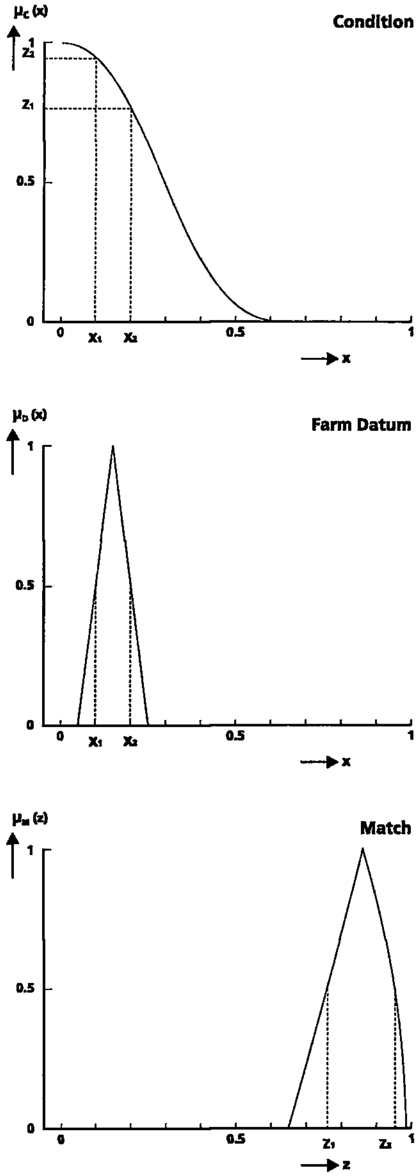


Figure 5.11 *MSFs for condition 'time of slurry application is at most bad', for farm datum 'bad', and for the resulting match. See text for explanation*

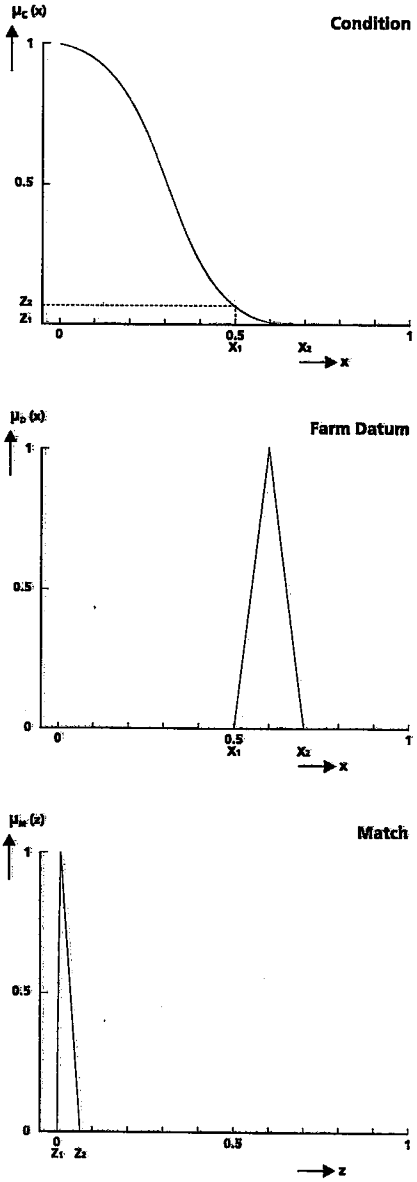


Figure 5.12 *MSFs for condition 'storage capacity of manure is at most low', for farm datum 'fair', and for the resulting match. See text for explanation*

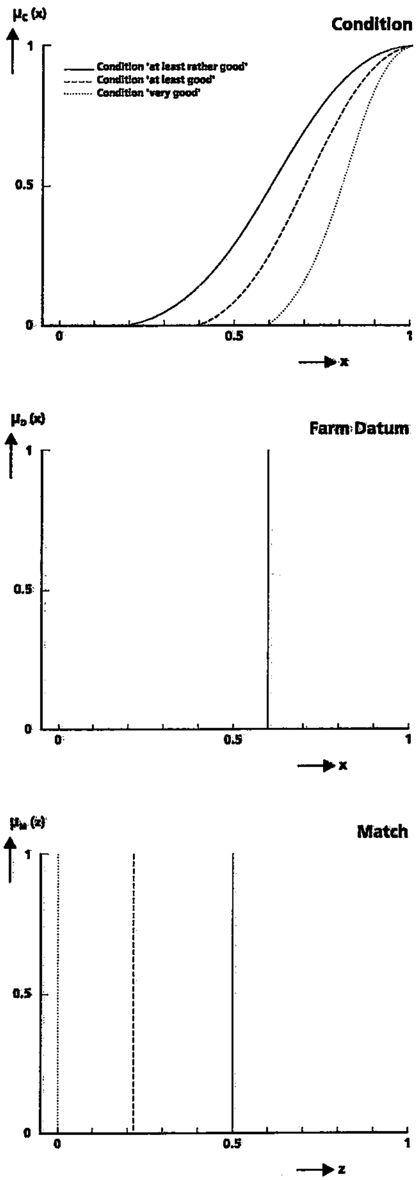


Figure 5.13 *MSFs for the conditions 'at least rather good', 'at least good' and 'very good', for the crisp farm datum 0.6, and for the resulting match. See text for explanation*

(0.05,0.25), resulting in an interval for the match of (0.65,0.99). From the MSF of the match it can be concluded that the match is good.

A very bad match is shown in figure 5.12, the third condition of our example in section 5.2.2. At x1=x2=0.6 the MSF value of the farm datum $\mu_D(x)$ is 1 in figure 5.12. At x=0.6 the MSF value for the condition $\mu_C(x)$ is 0. The membership value of the matching MSF $\mu_M(z)$ is 1 for z=0. For α-level 0.5 the interval [x1,x2] is [0.55,0.65], resulting in an interval for the match of [0,0.01]. For α-level 0 the interval [x1,x2] is (0.5,0.6), resulting in an interval for the match of (0,0.06). Figure 5.12 shows the procedure at this α-level.

How the matching MSFs in the figures 5.10, 5.11 and 5.12 of the three conditions of the example are combined with their weight (importance) to eventually infer the relevance of the conclusion is described in section 5.4.

Finally an example is presented where the farm datum is not a linguistic term but a numerical value. The farm datum for milk yield per cow has the value 7,400 kg on farm F. We have seen in section 5.2.3 that this value can be transformed to the value 0.6 of the ground set [0,1]. The MSF of this special case has already been shown in figure 5.4, and is identical to the figure in the middle of figure 5.13.

This farm datum of 0.6 (i.e. 'above average', crisp) can be matched with conditions. We take three different conditions as example: the milk yield per cow is 'at least rather good', 'at least good' or 'very good'. At x=0.6 the MSF value of the farm datum $\mu_D(x)$ is 1 in figure 5.13. At x=0.6 the MSF value for the condition 'at least rather good' is 0.5. Therefore the membership value of the matching MSF $\mu_M(z)$ is 1 for z=0.5. Since the farm datum has a crisp value, z=0.5 for all α-levels. The values of z for the conditions 'at least good' and 'very good' are 0.22 and 0 respectively.

Figure 5.14 shows (part of) the explanation facilities of FUZZY-DETECTOR concerning the matches between farm data and conditions for our example in section 5.2.2. The first bar, which represents the relevance of the conclusion according to the matching procedure is equal to figure 5.9. The next section explains the calculation of the relevance.
The second bar of figure 5.14 is derived from the MSF of the match in figure 5.10. The high uncertainty is evident. The last two bars stem from the figures 5.11 and 5.12.

Both relevance and uncertainty of the conclusion are mostly affected by conditions who are most important. In figure 5.14 it is shown that the relevance of the conclusion is somewhat uncertain, because (1) the farm datum for the first condition is uncertain and (2) the first condition is very important.

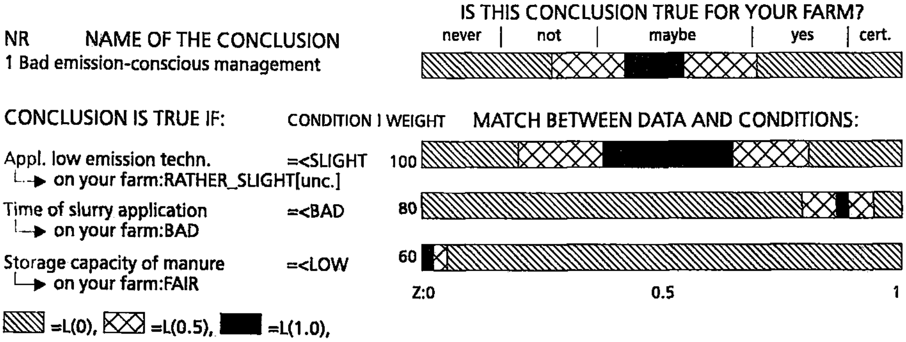


Figure 5.14 Explanation facility for the conclusion 'Bad emission-consious management'. Output from FUZZY-DETECTOR

The impact has been shown when the qualitative value of the farm datum for the first condition is not certain ([UNCERTAIN]). Sometimes it occurs that data are totally unclear or missing. FUZZY-DETECTOR can handle such cases in a very simple way.

Suppose that the farm datum from the first condition, 'Application low emission technique', was not clear or missing. The MSF for 'not clear' will then be used. The value of $\mu_D(x)$ is one for each value of x, because *if* the value of x *was* known without any uncertainty, *then* x could have been any value in the interval [0,1] with a maximum grade of membership of one. The MSF of the match is equal to the MSF of the farm datum in this case.

Figure 5.15 shows how the results from figure 5.14 are changed when the first condition is missing or not clear. Notice that the relevance of the conclusion is becoming considerably uncertain, also because the first condition is very important.

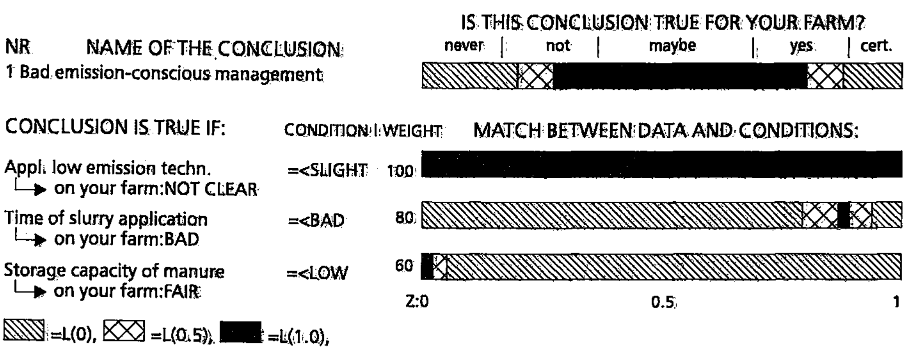


Figure 5.15 Explanation facility for the conclusion 'Bad emission-consious management' when the first farm datum is not known. Output from FUZZY-DETECTOR

Since the uncertainty of the relevance for this conclusion has increased, overlaps with other conclusions will be large. In this situation, it is for a decision-maker very difficult to chose among alternatives. So, decision making can be improved by additional information (i.e. replacement of 'not clear' with a linguistic or numeric value).

The value of information is different for each case. Generally, this value increases when the importance of a concept increases. When the third - moderately important - concept was not clear instead of the first, the uncertainty of the final conclusion would be less.

5.4 Relevance of the conclusion

The presentation of the relevance of the conclusion with FUZZY-DETECTOR has been illustrated earlier in this chapter (figure 5.9). In this section the procedure and algorithms are described for calculating the relevance.

The relevance of a conclusion stated in the THEN part of an IF-THEN rule is a function of each condition with its importance as stated in the IF part of the rule, and each farm datum with its certainty status. It should have the property that the better the farm data match important conditions the higher the value of the relevance is. In section 5.2.3, we defined a relevance function r with this property. The first step in the definition of this function is the calculation of the value m_i (i.e. the values $\mu_M(z)$) of the matching function defined for condition C_i, $i=1,...,n$, and the farm datum D_i, $i=1,..,n$, where n is the number of conditions in the rule. This was done in section 5.3. The second step is the definition of a weight w_i reflecting the importance of conditions C_i. The relevance function r is defined as the weighted average of the values m_i with weight w_i, $i=1,...,n$. The algorithm for the construction of this function is explicated below. Note that this function has the desired property discussed above.

According to the fuzzy set theory, the membership function μ_r of the relevance function r is defined by $\mu_r(z)$, the supremum over all weighted averages defined by m_i, w_i, $i=1,..,n$, which are equal to z or the minimum of $\{\mu_M(m_i),\mu_W(w_i)|i=1,...,n\}$. As with the matching function the MSF can be expressed by its level sets $L(\alpha)$, $0\leq\alpha\leq1$. These level sets can be calculated by the algorithm of Kwakernaak (1979). We describe the algorithm below and refer to the original paper for the correctness proof.

Algorithm 1) for the level set L(α) of r (Kwakernaak, 1979)

Step 1. Define the numbers

$m1_i=\min\{m_i|\mu_M(m_i)\geq\alpha\},\quad i=1,...,n.$

$m2_i=\max\{m_i|\mu_M(m_i)\geq\alpha\},\quad i=1,...,n.$

$w1_i=\min\{w_i|\mu_W(w_i)\geq\alpha\},\quad i=1,...,n.$

$w2_i=\max\{w_i|\mu_W(w_i)\geq\alpha\},\quad i=1,...,n.$

Step 2. Sort the values $m1_i$, $i=1,...,n$. Renumber the conditions such that $m1_1\leq m1_2\leq...\leq m1_n$. Calculate:

$$a = \min_{0\leq j\leq n} \frac{\sum\limits_{i=1} w2_i m1_i \quad + \quad \sum\limits_{i=j+1} w1_i m1_i}{\sum\limits_{i=1} w2_i \quad + \quad \sum\limits_{i=j+1} w1_i}$$

Step 3. Sort the values $m2_i$, $i=1,...,n$. Renumber the conditions such that $m2_1\leq m2_2\leq..\leq m2_n$. Calculate:

$$b = \max_{0\leq j\leq n} \frac{\sum\limits_{i=1} w1_i m2_i \quad + \quad \sum\limits_{i=j+1} w2_i m2_i}{\sum\limits_{i=1} w1_i \quad + \quad \sum\limits_{i=j+1} w2_i}$$

Step 4. The level set L(α) is given by the interval [a,b].

As explained in section 5.2.3, only the level sets L(0.5) and L(1) of the relevance function are calculated and displaid by the computer. Since L(0) (i.e. $\alpha\geq 0$) always yields the ground set [0,1], this α-level is not calculated.

Let us apply the algorithm to calculate the level sets L(0.5) and L(1) for the conclusion 'Bad emission-conscious management' for the rule and farm data defined in section 5.2.3. The matching MSFs have been presented in section 5.3 (see figures 5.10, 5.11 and 5.12). The MSFs of weights are given in section 5.2.2 (figure 5.7). Table 5.1 lists the values required in the algorithm to calculate the level sets L(0.5) and L(1).

---

1) The algorithm is based on the method of weighted summation. The final rating R_i of a particular alternative i can be calculated from the ratings r_{ij} of criteria j and their weighting coefficients (or importance) w_{ij} by $R_i = \Sigma(r_{ij}w_{ij}) / \Sigma(w_{ij})$ (Janssen, 1991). This algorithm can only be used for numbers, and not for intervals from fuzzy sets.

Table 5.1 Values for $m1_i$, $m2_i$, $w1_i$ and $w2_i$ to calculate the values for a and b at the level sets L(0.5) and L(1)

	$m1_1$	$m1_2$	$m1_3$	$m2_1$	$m2_2$	$m2_3$	$w1_1$	$w1_2$	$w1_3$	$w2_1$	$w2_2$	$w2_3$	a	b
L(0.5)	0.17	0.78	0.00	0.83	0.94	0.01	0.9	0.7	0.5	1.0	0.9	0.7	0.298	0.700
L(1.0)	0.35	0.88	0.00	0.65	0.88	0.00	1.0	0.8	0.6	1.0	0.8	0.6	0.439	0.564

The MSF of the relevance has already been shown in figure 5.8.

5.5 The use of FUZZY-DETECTOR in KBSs

The tool FUZZY-DETECTOR in which the method has been implemented can be applied to build KBSs. The conclusion 'Bad emission-conscious management' has been used as an example in the previous sections. Another rule 1) from the same domain may look like:

IF
(Bad emission-conscious management is
>RATHER_RELEVANT) [VERY_IMPORTANT]
AND
(Application of Nitrogen-fertiliser is >HIGH) [VERY_IMPORTANT]
AND
(Stocking rate is >HIGH) [IMPORTANT]
AND
(General impression hygienic condition
is <BAD) [MODERATELY_IMPORTANT]
THEN
Bad utilisation animal manure

The relevances of conclusions can be used as conditional concepts in other rules. As shown in this rule, the conclusion from the rule in section 5.2.2, 'Bad emission-conscious management', can be matched with the condition >RATHER_RELEVANT. Successive calculations follow the same algorithm.

This process is analogous to forward chaining in the literature about artificial intelligence (e.g. Winston, 1984). Rules produce facts (conclusions), which may be used in other rules to produce new facts, and so on.

1) The knowledge for this rule is also supplied by ing. H.H. Luesink from LEI-DLO.

The two rules presented so far could stem from the domain called 'Efficient nutrient management'. Figure 5.16 shows the relevances of some conclusions in this domain in sorted order.

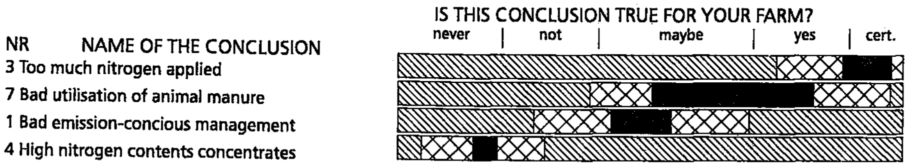


Figure 5.16 Relevances of conclusions from the domain 'Efficient nutrient management'. Output from FUZZY-DETECTOR

The explanation facilities, which the user can ask for in the tool FUZZY-DETECTOR, were already shown in figure 5.14.

5.6 Expert's role in FUZZY-DETECTOR

The expert's role is limited to the supplier of knowledge for the construction of the rule base (e.g. the rule in section 5.2.2). The process of knowledge acquisition is comparable with the method IMAGINE, even quicker and easier. After the expert has been asked to concentrate on a certain conclusion, all he has to do is to name the concepts, conditions, importances, and the interval for the transformation function (see section 5.2.1). The whole process is called *backward knowledge acquisition*, in accordance with the used term in chapter 4 for IMAGINE.

After rules are stored in the rule base, FUZZY-DETECTOR automatically matches farm data against all rules to infer the relevance of *all* conclusions.

It might be possible to let the expert himself construct the MSF for the valuation of a certain concept regarding the conclusion, as alternative for the MSF of the matching (result of the matching algorithm). This would increase the reliability, but may not be workable. It is too abstract and time consuming for the expert to do. The expert's presence in that situation is necessary to judge each case (or farm) to construct the alternative for the matching graph. It is also doubtful whether a high accuracy in the function's construction is in agreement with the applied, rather rude, method FUZZY-DETECTOR and its algorithm.

In the example used, it is assumed that the expert does not doubt the accuracy of the rule in the knowledge base. His doubt, however, can be made explicit in FUZZY-DETECTOR when the rules are defined. When this is the case, the intervals of the final MSF shall become wider, depending on the expert's doubt or uncertainty. This is somewhat comparable with the rule's certainty factors applied in traditional rule-based systems (e.g. Waterman, 1986).

Not shown in the example of this chapter is the possibility that the expert denotes additional uncertainties for the conditions and importancies. The linguistic expression *important[uncertain]*, for example, is treated comparably to the farm datum in figure 5.3.

5.7 Concluding remarks

LEI-DLO developed the tool FUZZY-DETECTOR for building KBSs in domains where data might be qualitative, uncertain and incomplete. A tendency in the development of agricultural software is to take into account the farmer's individual goals, preferences, skill, capacity and style of farming. There are certainly situations where the variables can no longer be expressed numerically but only in linguistic terms. These are typical problem situations for the fuzzy set theory (Kickert, 1978). The presented method in FUZZY-DETECTOR is an attempt in this direction.

Most traditional KBSs are rooted in a two-valued logic, and thus the rules must be executed in an all-or-nothing manner (Whalen and Scott, 1983). The conclusion or action is only true when the whole condition set is true. In FUZZY-DETECTOR *each* conclusion is more or less true, represented by intervals from its MSF. With this tool, it is possible to develop rule-based systems where the rules are in fact 'fuzzy IF..THEN rules'. Although no information is lost, large systems built with FUZZY-DETECTOR might be time consuming during consultation since all conclusions are tested. Introduction of crisp or 'hard' conditions in the fuzzy rules and the creation of a structure with rule sets can manage the problem.

The use of the fuzzy set theory in KBSs is also defensible from the expert's point of view. 'Since the knowledge base of an expert system is a repository of human knowledge, and since much of human knowledge is imprecise in nature, it is usually the case that the knowledge base of an expert system is a collection of rules and facts which, for the most part, are neither totally certain nor totally consistent.' (Zadeh, 1983). In short, for both expert and user, the fuzzy set approach may be characterised as a humanly perceived approach (Nagaki, 1992).

The most important aspect of FUZZY-DETECTOR is the management of uncertainty concerning both expert's knowledge and the data. In many KBSs, uncertainty is expressed in certainty factors (e.g. Waterman, 1985). The computation of certainty factors is based on two-valued logic and probability theory. According to Zadeh (1983) this is an invalid way, suggesting that certainty factors must be represented as fuzzy rather than crisp numbers.

In the way it is presented in this article, the method FUZZY-DETECTOR must not be seen as a rigid one. It can easily adjust to specific demands concerning a particular domain. A most important issue will be a possible redefinition of MSFs. These are just subjective evaluations and, consequently, all problems arising with fuzzy sets are due to the lack of our knowledge of the interpretation of 'fuzzy' by such functions (Dombi, 1990). So, although MSFs are the very core of the fuzzy set theory (Negoita, 1985), it is not surprisingly that they are often criticised (e.g. by French, 1984).

At the moment, I have proposed a number of MSFs applied in FUZZY-DETECTOR and these are predominantly based on my subjective opinion supplied by findings from the literature (e.g. from Baas and Kwakernaak, 1977). Since they are not validated, much attention should be paid to such functions in the future.

For agriculture, and agricultural economics research in particular, the application of the fuzzy set theory may introduce a number of opportunities, especially in combination with current traditional methods. There are possibilities in the areas of optimisation (LP), prediction and forecasting of events, monitoring, interpretation from numerous data, management support for farmers, and marketing. 'Much of the decision making in the real world takes place in an environment in which the goals, the constraints and the consequences of possible actions are not known precisely.' (Bellman and Zadeh, 1970).

6. GLOBAL-DETECTOR: KNOWLEDGE-BASED SYSTEM FOR ANALYSIS AND DIAGNOSIS OF PERFORMANCE ON DAIRY FARMS

"Ik ben heel nieuwsgierig en wil zoveel mogelijk weten. Met elk advies en elke analyse kan ik mijn voordeel doen." 1)
(Dairy farmer B. Prins about GLOBAL-DETECTOR, 'Agr. Dagblad', 7(1993)228:2)

'The best way to ensure acceptance of a system is to be very complete in preliminary interviews with potential users (and designing the system in accordance with users needs and goals) and to build a restricted prototype and have users give responses to questions.'
(Gordon et al., 1987)

In this chapter, GLOBAL-DETECTOR will be presented. This is a knowledge-based system (KBS) for the global analysis of year-end results (from farm accounts) concerning aspects of gross margin from dairy farms. GLOBAL-DETECTOR is developed according to the requirements described in section 1.4 as much as possible. The requirements of GLOBAL-DETECTOR (and ENVIRONMENT-DETECTOR, chapter 7) are:

- use of already available data;
- support tactical decision-making and give suggestions for improvement;
- take into account farmer's specific situation;
- give much insight;
- stimulate farmer's creativity;
- easy and fast maintenance;
- advocate widespread use by individual farmers and extension services of different organisations.

The system tries to fill the gaps of the lack of good performance figures and the lack of good farm comparison. With this instrument, the analysis of accounting data by farmers may be improved.

The methods FAS (farm-adjusted standard), IMAGINE and FUZZY-DETECTOR as described in earlier chapters, are applied in GLOBAL-DETECTOR. FASs are used to position farm results with respect to results of comparable farms. The deviations, which are clues for good or bad management, are analysed by the Artificial Intelligence tools IMAGINE and FUZZY-DETECTOR. The user may choose one of these two tools. The result of the analysis is a list of strong and weak aspects regarding the

---

1) "I am very inquisitive and want to know as much as possible. I can take advantage with each advice and every analysis (transl.WH)".

farm and farm management as well as suggestions for improvement. Both tools perform the same task, by analysing the same knowledge base.

The objective in this chapter is for the most part the technical and conceptual description of GLOBAL-DETECTOR, taken into account that the system has to fit the above-mentioned requirements as good as possible. It might throw light on the possibilities for the development of such a KBS for the analysis of technical and economic data from individual dairy farms to support the tactical management based on the expert's knowledge. The illustrations in this chapter are restricted to returns and variable costs. An analysis of fixed costs is possible with the system, although no knowledge base has been developed yet for a diagnosis of these costs.

At the end of this chapter, special attention is given to accounting for user's need and management behaviour, to the evaluation of GLOBAL-DETECTOR and user's experience with it, and to the different ways the system might be used.

6.1 Some technical specifications of GLOBAL-DETECTOR

GLOBAL-DETECTOR has been developed from scratch by means of an Artificial Intelligence's language. This language, muLISP (Soft Warehouse), is a dialect of the standard language Common LISP (Steele, 1984). Only a small amount of memory is consumed by muLISP, which is also relatively fast. Software is developed in muLISP for all functions of GLOBAL-DETECTOR, i.e. software for user interface, for calculation of FAS, for making tables, graphs and bar diagrams, as well as explanation facilities, and for the application of IMAGINE and FUZZY-DETECTOR. No additional software packages are used.

All software for control, inferences, graphical output, etc, have been programmed domain-independently, which means that this software can be used as a 'tool' or 'shell' for developing similar systems in other domains, also outside agriculture.

GLOBAL-DETECTOR can be consulted on an IBM PC or compatible computer. A hard disk is recommendable. About 400 kByte of internal- and about 350 kByte of external memory is sufficient. These modest requirements make it possible to use this system on farms.

After the user (farmer, advisor) has started GLOBAL-DETECTOR for an analysis, accounting data from a chosen farm in a chosen year are read in from a disk, which is followed up by the calculation of FAS values and other relevant variables. Subsequently series of possibilities appear on the display which can be used to select the specific information the user wishes to go into just by typing the number of interest. This menu-structure appears to be very user-friendly.

6.2 The farm data

GLOBAL-DETECTOR requires about fifty data from farm accounts for global analysis. These farm data, which are already available, can be subdivided in the following types:

- general data (e.g. number of cows, area of land);
- returns (e.g. cattle credits);
- variable costs (e.g. feeding costs, veterinary costs);
- fixed costs (e.g. costs for labour, buildings);
- production data (e.g. amount of fertilizer, milk yield);
- performance data (e.g. gross margin).

A user of the system is asked for the farm number and the desired year of analysis. The required farm data are not obtained by direct access from the data base of LEI-DLO but from intermediate data files. This makes it possible to use the system on farms. A farmer who uses GLOBAL-DETECTOR on his own farm has (small) data files with his own data in them. He has no access to the data of other farms. The regression coefficients for the FAS models and the standard prices for the year of analysis are stored in the internal memory.

GLOBAL-DETECTOR may be used by other accountancies outside LEI-DLO. The problem of uniformity (chapter 1) can be faced by the development of data transformation programmes and by a number of necessary adjustments of GLOBAL-DETECTOR. This has already been performed successfully for two accountancies.

6.3 Analysis of the farm data

Analysis is necessary for providing insight into the strong and weak parts of the farm (Dobbins, 1989) and is therefore the focal point of any record-keeping activity (James and Stoneberg, 1986). In earlier chapters (1 and 2) the need for good performance figures for reference is stressed and it has been concluded that the use of FASs is satisfactory for application in the KBSs, GLOBAL-DETECTOR included.

6.3.1 The farm-adjusted standards (FASs)

The method of the FAS is a new method of external farm comparison. This method has been developed because there were problems in comparing an individual farm with good comparable farms. This method tries to tackle that problem. The FAS and the way in which FAS models have to be developed was explained in chapter 2. The FAS models in GLOBAL-DETECTOR must be regarded as the core of the system, as will be made clear in the rest of this chapter.

FAS models that are used in GLOBAL-DETECTOR are developed by De Haan (1991) for most returns and variable and fixed costs. Each return and cost factor has its own specific FAS model or, stated otherwise, a specific set of independent variables with or without some of their interactions.

When a particular farm F in year Q is analysed, FAS values are calculated by means of the year-specific FAS models. Figure 6.1 shows the output from GLOBAL-DETECTOR for farm F in year Q as example. All aspects are expressed in the same reference: NLG per hectare. This is a justifiable point of view, since it might be expected that the milk quota per hectare will not change on the short term on one particular farm, and farmers aim at milking their full quota at the lowest possible costs (De Hoop et al., 1988). The milk quota per hectare is thus an important production constraint on a Dutch dairy farm.

DETECTOR 11111/92

FARM RESULTS IN NLG PER HECTARE	RESULT	STANDARD	DEVIATION	*/!
Gross margin	7,339	7,771	-432	!
YIELDS				
Milk receipts	9,365	9,344	21	*
Cattle credits	934	1,177	-243	!
Remaining	202	196	6	
DIRECT COSTS				
Additional feeding	1,635	1,432	203	!
Veterinary	194	229	-35	*
Insemination	134	117	17	!
Milk recording+Herdbook	55	63	-8	
Interest	337	332	5	
Milk products	114	110	4	
Contract rearing	0	0	0	
Other cattle costs	143	95	48	!
Seeds+Chemicals	31	33	-2	
N-fertiliser	466	466	0	
Other fertilisers	34	41	-7	
Other costs crops	0	2	-2	
Minerals, etc	19	26	-7	

Do you want some background information about the figures? Y/N _

Figure 6.1 Realised values, standard values (calculated with FAS models) and deviations between realised and standard values for gross margin, returns and variable costs on farm F with number 11111 for year 1992/93. Output from GLOBAL-DETECTOR

For each return and (variable) cost factor, a FAS value has been calculated just by putting the values of the (independent) variables in the models. This will be illustrated for cattle credits. The following model is used to calculate the FAS value for cattle credits per cow (see also formula 2.2 in chapter 2):

$$FAS_{cc\text{-}av} = c_{av,Q} + \beta_{1av,Q} * FPCM + \beta_{2av,Q} * EGVEORMK + \beta_{3av,Q} * BREED \quad (6.1)$$

The year-dependent coefficients ($c_{av,Q}$, $\beta_{1av,Q}$, $\beta_{2av,Q}$, $\beta_{3av,Q}$) of the FAS model were estimated earlier and entered in the system GLOBAL-DETECTOR. When this farm F is analysed, the values for (corrected) milk yield ($FPCM_F$=7,560), number of cattle per cow ($EGVEORMK_F$=0.392) and the code for the breed ($BREED_F$=92) are placed in the model. The calculated FAS value for cattle credits per cow is then 731 NLG. Since all data in figure 6.1 are expressed in NLG per hectare, this value per cow is multiplied by the number of cows per hectare on this farm (1.61) to obtain the FAS value 731*1.61=1,177 (second column of data in figure 6.1). Our farm F has realised a value for cattle credits per hectare of only 934 (first column of data in figure 6.1), while the average Dutch farms in the same year with the same milk yield, number of animals and breed had a value of 1,177. This deviation (-243) is presented in the third column of figure 6.1. The symbol * in the last column denotes a favourable deviation, while the symbol ! denotes an unfavourable one. These deviations are important information sources for the expert to conclude strong and weak aspects and suggestions for improvement (see later in this chapter).

No FAS models have been developed for contract rearing and nitrogen fertiliser. The majority of farms do not apply contract rearing and those who do, have high costs for this aspect. A FAS is therefore meaningless for contract rearing. The amount of nitrogen fertiliser is used as an independent variable in the FAS models because this variable has much influence on the grassland production and from that on the amount of additional feed purchased. For a good comparison of the additional feed costs we have decided that a FAS model for the costs of nitrogen fertiliser needs not be developed.

The FAS model of gross margin per hectare is composed of the individual FAS models for the various returns and variable costs. These FAS models are based on the purchase price of feed. The FAS values for gross margin and additional feeding costs per hectare are therefore corrected for a price difference between selling and buying of roughage when the farm is in a position that it should have a surplus of roughage (the FAS for purchase of roughage is negative then).

The user of GLOBAL-DETECTOR can get on-line information with respect to the FAS models used, and to see which aspects of the farmer's specific situation is accounted for.

After the user has analysed his results in a particular year, he can easily skip to another year for analysis. Data from that year are automatically read in from the data base, and regression coefficients and standard prices are read from the internal memory. The calculations that follow are done by means of FAS and standard prices for that particular year (see section 6.3.3). As far as a very recent year is concerned, and having no equations for FAS available at the moment, the most recent equations are used in combination with price indices to correct partly for year influences (chapter 2).

6.3.2 Graphical presentation of results and effects

The FAS models are used for graphical presentations. The user might be interested in how an aspect is affected by an independent variable. A few dozens of graphs are at his disposal. One example is the influence of the milk yield (corrected for the percentage of fat and protein) on the gross margin per hectare at the specific level of milk quota per hectare (figure 6.2). The value of farm F is denoted by a small block, the FAS values for different levels of milk yield are expressed in the curve. All other independent variables are kept constant. The milk quota per hectare is also kept constant, which means that the number of cows per hectare has to decrease when the milk yield increases. Although such a graph must be interpreted with great care (chapter 2), it is obvious

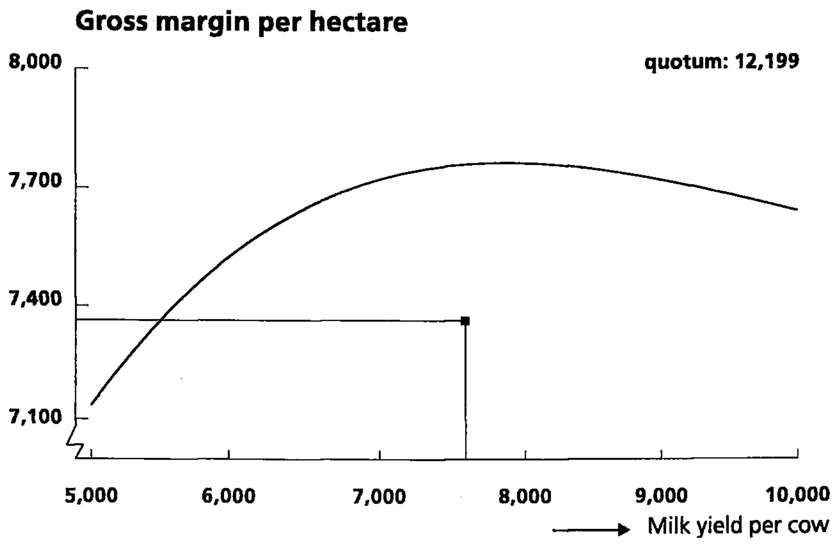


Figure 6.2 Relation, based on the FAS models, between the milk yield per cow (X-axis) and the gross margin per hectare (Y-axis) at the same milk quota per hectare as farm F (12,199); and the actual position of farm F (represented with a small block) for the year 1992/93. Output from GLOBAL-DETECTOR

from figure 6.2 that for this particular farm F the gross margin per hectare will not always increase when the milk yield increases. Daatselaar (1988) came to comparable conclusions in his research.

Some graphs do not only display the lines based on all farms (like in figure 6.2), but also lines for the highest and the lowest performing 25% of the farms for that aspect. Figure 6.3 displays the influence of the corrected milk yield per cow on the cattle credits per cow. The upper line shows the FAS for the highest 25% of the farms corrected for the same independent variables. One might call these farms 'best 25%', but this judgement may not be used when only one separate aspect is highlighted as in this case. It is only permitted when the farm as a whole is judged.

The value on farm F is lower than that of a comparable average farm, but higher than the average of the lowest quarter of farms (lowest line in figure 6.3). Reaching the average standard may be a goal for farmer F.

A great number of such relations can be shown to the user when he wants to. Another form of graphical presentation is shown in figure 6.4 as an example. The farmer's position with regard to the buying of additional feed (expressed in net energy units) is distinguished for the types

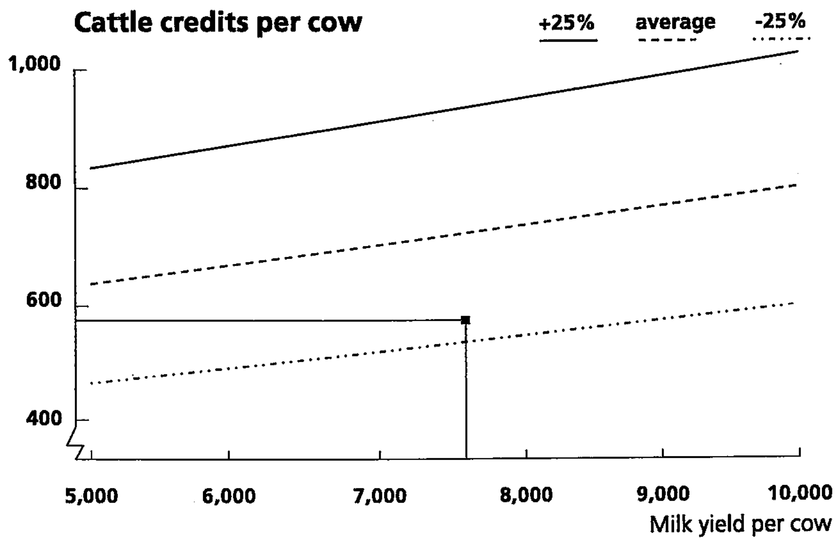


Figure 6.3 Relations between the milk yield per cow (X-axis) and the cattle credits per cow (Y-axis) for year 1992/93; and the position of farm F. Average, +25% and -25% relation is based on FAS models for all farms, the 25% 'best' and the 25% 'worst' farms regarding cattle credits per cow. Output from GLOBAL-DETECTOR

of feedstuffs: concentrates, fibrous roughage and roughage without fibers. The total amount of purchased feed is displayed by the bars on the left of figure 6.4. The little block is again the position of farm F. The upper side of each bar indicates the FAS value. The lowest bar for the 25% of farms with the lowest amount, the bar in the middle for the average and the highest bar for the 25% of farms with the highest amount. As can be seen, the total amount of purchased feed on farm F is higher than an average comparable farm (with the same intensity, etc), but lower than comparable farms with the highest amount of purchased feed. The other groups of bars show, from left to right, farmer's position with respect to the amount of concentrates, the total amount of roughage, the amount of fibrous roughage and the amount of roughage without fibers.

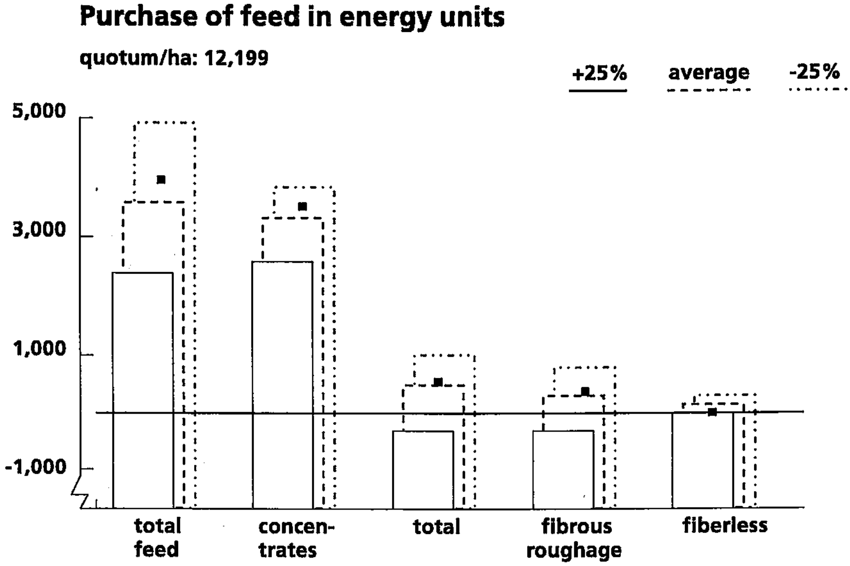


Figure 6.4 The farmer's position (little block) with respect to standards (FAS values for average, lowest 25% and highest 25%) for different types of purchased feed stuffs in year 1992/93 (see text for explanation). Output from GLOBAL-DETECTOR

All graphical presentations are joined up with the relevant explanation. The user can read the text to understand the meaning of the graph, in order to give him the required insight. Relations are made visible, GLOBAL-DETECTOR is therefore no black-box. Differences in input-output relations between different groups of farmers (+25% and -25%) are shown and the farmer can use his creativity to get an idea about his own input-output relation based on his position and the input-output relations shown. In chapter 2 it has been mentioned that individual input-output relations are not known. Graphs like the ones shown above might stimulate a farmer's creativity in finding his own relation.

Besides a great number of graphs, there are also a handful of tables, e.g. the estimated effects of selling or buying quota or land. Figure 6.5 shows the estimated effect of a decrease and an increase of quota per hectare on gross margin. Quota costs are excluded in the expected gross margin, and all other variables (except number of livestock per hectare) remain the same. The expected outcome is accounted for the farm specific situation and specific input-output relation (chapter 2). The expected outcome gives the user an idea of the direction and magnitude of the effects.

DETECTOR 11111/92

EFFECT OF A CHANGE IN QUOTA PER HECTARE ON GROSS MARGIN

DECREASE/INCREASE QUOTA	NEW QUOTA	GROSS MARGIN REALISED	GROSS MARGIN EXPECTED
-1,500	10,699		6,561
-1,000	11,199		6,848
-500	11,699		7,094
0	12,199	7,339	7,339
500	12,699		7,586
1,000	13,199		7,833
1,500	13,699		8,079

Remark: The expected gross margin is calculated from an average relation. On your farm, the effect might be slightly different.

Press SPACEBAR... _

Figure 6.5 The expected effect of a change in quota per hectare on the expected gross margin per hectare. Output from GLOBAL-DETECTOR

6.3.3 Analysis of historical data

The user of GLOBAL-DETECTOR can analyse his results of three to five successive years. Figure 6.6 presents a three-year analysis for farm F. This figure is comparable with figure 6.1, the FAS values are not shown because of space limitations.

With such a figure, the farmer gets useful information about how his position has changed with respect to comparable farms for the different returns and variable cost aspects. The gross margin per hectare was unfavourable in these three years and the cattle credits developed unfavourably. Since this farm is compared with similar other farms in the *same* year, year-effects have to be excluded in the interpretation.

GLOBAL-DETECTOR additionally presents a summary of the major levels and trends. Some simple heuristics are implemented to make a

DETECTOR 11111/92

RESULTS FROM 3 YEARS (NLG/HA)	RES *90*	DEV	RES *91*	DEV	RES *92*	DEV
Gross margin	8,920	-228	7,431	-272	7,339	-432
YIELDS						
Milk receipts	10,703	-92	9,275	62	9,365	21
Cattle credits	1,629	-2	934	-236	934	-242
Remaining	226	-60	166	8	202	6
DIRECT COSTS						
Additional feeding	1,974	129	1,294	19	1,635	203
Veterinary	223	-22	192	-32	194	-35
Insemination	152	7	235	115	134	17
Milk recording+Herdbook	46	-33	82	13	55	-8
Interest	426	16	366	-18	337	5
Milk products	129	-26	106	-8	114	4
Contract rearing	0	0	0	0	0	0
Other cattle costs	147	43	165	72	143	48
Seeds+Chemicals	37	3	153	115	31	-2
N-fertiliser	465	0	435	0	466	0
Other fertilisers	39	-23	36	-27	34	-7
Other costs crops	0	-1	0	-1	0	-2
Minerals, etc	8	-11	31	8	19	-7

Do you want some background information about the figures? Y/N _

Figure 6.6 Realised values and deviations from standards (based on FAS models) for gross margin, returns and variable cost components of three successive years. Output GLOBAL-DETECTOR

trend analysis of the data. The user can also ask for graphical information, which shows not only the levels of results and FAS values, but also the development of other variables, like the milk yield, the prices, etc.

The historical analysis also gives a summary of the strong and weak aspects and suggestions for improvement which were valid in the successive years, together with an overview of shifts during the years. The way these aspects and suggestions are deduced is explained in section 6.4.

6.3.4 Most striking features

The user can ask for a display of the most striking features on the farm, combined with the way GLOBAL-DETECTOR has inferred these. The function of this overview is to give the farmer or advisor a quick idea of some outranging data in order to pin-point them on the account at hand. This may be important for the identification of problems.

Algorithms for inferring the most striking features are derived from both descriptive statistics and plain heuristics from an expert. The displayed features are not strong or weak aspects, they are merely characteristics worth mentioning when an expert takes a quick glance at the account. Strong and weak aspects of the farm management and suggestions for improvement are the result of the KBS in the diagnosis part of GLOBAL-DETECTOR.

6.4 Diagnosis of the farm

Unfavourable deviations from FASs do not necessarily imply weak aspects. For example, a high cost factor may result in a high return factor. Knowledge or expertise is indispensable for evaluating deviations in combination with other factors to make a sound diagnosis of the performance. Dobbins (1989) emphasises the role of an 'expert' for this task. For GLOBAL-DETECTOR one person was eligible for the role of expert (De Hoop from LEI-DLO). He has not only expertise in the domain, he was also acquainted with the FASs and the method IMAGINE.

The objective of diagnosis is to find out what is wrong in the economic and/or technical situation of the farm (Longchamp et al., 1990) in order to provide the manager with information allowing the performance to be improved (Dobbins, 1989).

Still being in GLOBAL-DETECTOR, the user may skip from the analysis part to the diagnosis part, simply by choosing from the menu. There are two ways for inferring suggestions: by IMAGINE and by FUZZY-DETECTOR. The user may choose among these two methods. Knowledge bases are comparable for both methods with respect to GLOBAL-DETECTOR. IMAGINE is more accurate but less understandable than FUZZY-DETECTOR. Although FUZZY-DETECTOR has the possibility to use qualitative and uncertain data, only numerical data are used in GLOBAL-DETECTOR. Section 6.6 goes into the evaluation and validation of the diagnosis with IMAGINE and FUZZY-DETECTOR.

6.4.1 The method IMAGINE for knowledge acquisition and representation

IMAGINE is described in chapter 4. The development of the KBS GLOBAL-DETECTOR led to the development of the method IMAGINE, because reasoning with quantitative data was necessary.

A separation between strong aspects, weak aspects and suggestions for improvement will not be made. Instead, they will be taken together and called conclusions from now on. In total there are two dozens of different conclusions and each conclusion will be treated in the same way.

The relevant conclusions or performance judgements appear on the screen (figure 6.7). Like on most farms, there are only a handful of conclusions that are relevant. In this example there are three conclusions relevant, which happen to be three suggestions for improvement of the income.

Each conclusion has a certainty, expressed in a certainty factor. This factor ranges from -100 (absolutely not true) to +100 (absolutely true). A value of 0 indicates indifference. Only conclusions with a certainty factor above +20 are presented in figure 6.7. GLOBAL-DETECTOR shows additionally on a following screen an overview of indifferent conclusions (certainty factor between -20 and +20) and an overview of conclusions that are not relevant (certainty factor below -20).

DETECTOR 11111/92

NR.	RELEVANT CONCLUSIONS FOR YOUR FARM	RELEVANCE
12	Improve your feed and grassland management	82
6	Decrease the amount of concentrates/cow, milk yield may decrease	58
1	Decrease the manuring with nitrogen	36

I=info about NR G=explanation of RELEVANCE H=help C=continue

Figure 6.7 Most important conclusions for farm F inferred by the method IMAGINE. Output from GLOBAL-DETECTOR

The user can retrieve extensive information about the meaning of the certainty factors and how the expert has reached a certain conclusion by means of the explanation facilities of GLOBAL-DETECTOR. The explanation might stimulate the creativity of the farmer since he can try to find out if the suggestions for improvement and the way they are inferred fit his own situation and thoughts. After he has 'digested' this information, the creativity can be stimulated again since he may want to find ways that can improve his situation (e.g. make a detailed analysis of his own or visit an advisor and ask him directed and relevant questions).

The following displays are shown to the user for the necessary information:

- an easy readable text about how the expert in general comes to such a conclusion and why this conclusion is or is not relevant for the analysed farm. This serves as background information;
- a handful of options for graphical information (see section 6.3.2.) about the conclusion at hand;
- the bar graphs from the IMAGINE method. As an example, figure 6.8 shows how the system comes to the conclusion 'Lower the amount of concentrates per cow; milk yield may decrease'. The user which is not so familiar with the interpretation of that information can ask for help from GLOBAL-DETECTOR.

Figure 6.8 will only be explained shortly here. Five information sources (data or variables, e.g. milk yield per cow) are needed to infer the conclusion. The conclusion is supported by values on the right side of the bars. For example, a milk quota per hectare below 11,000 and a milk yield above 6,500 both support the conclusion. Counteracting values appear on the left side of the bars, like a milk quota per hectare above

11,000. In our example the conclusion is counteracted by a high milk quota per hectare (Δ=12,199), resulting in a negative individual score of -1.2. Since all other variables support the conclusion, the average score has a positive value of 2.34 (by compensation). This value is transformed to a relevance of 58 or 'rather relevant' (figure 6.7). See chapter 4 for additional detailed information on the method IMAGINE.

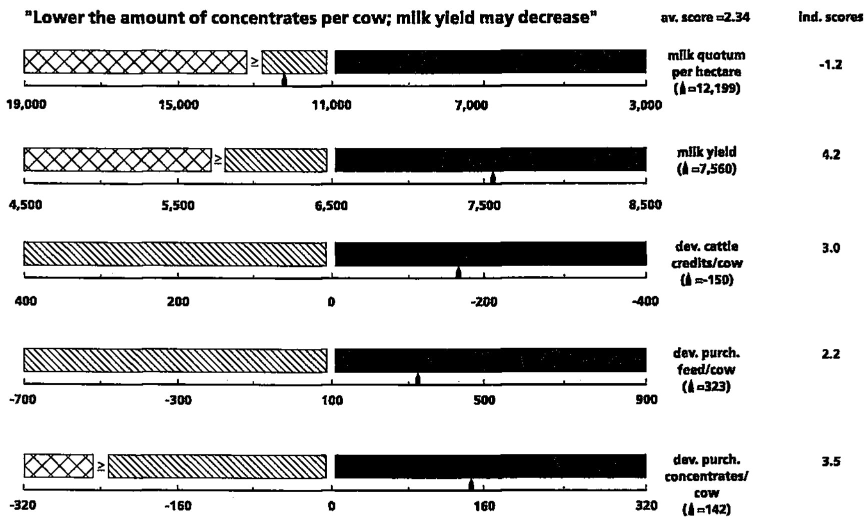


Figure 6.8 Detailed explanation for the conclusion 'Lower the...' of farm F, by IMAGINE. Output from GLOBAL-DETECTOR. (AV=average; dev. means deviation between realised and FAS value)

Information as presented in figure 6.8 is not very clear for most farmers, despite help from GLOBAL-DETECTOR for interpretation. Farmers who tested GLOBAL-DETECTOR asked for a simpler method. This led to the use of the method and tool FUZZY-DETECTOR in the system GLOBAL-DETECTOR.

6.4.2 The method FUZZY-DETECTOR as an alternative for IMAGINE

The knowledge base of IMAGINE has been rewritten by the knowledge engineer to produce a knowledge base for the tool FUZZY-DETECTOR. This could be done within a few hours, independently from the expert. Both knowledge bases are comparable and contain the same knowledge and conclusions.

After some tests were performed, the results showed unacceptable deviations with the results by the method IMAGINE. Most problems arose from the fact that IMAGINE contained rejection values, while FUZZY-

DETECTOR did not. A rejection value in IMAGINE counteracts the conclusion increasingly with increasing (or decreasing) values of a variable, and makes compensation in some situations intendedly impossible (chapter 4). In FUZZY-DETECTOR, such unintended compensations do occur however.

To solve the problem, the possibility of (a smooth) rejection is also introduced in the tool FUZZY-DETECTOR to model the knowledge for GLOBAL-DETECTOR. This pragmatic solution has not been presented in chapter 5.
The method FUZZY-DETECTOR assigns to each conclusion i a degree of relevance, $R_i \varepsilon [0,1]$. If one or more conditions j of conclusion i have a rejection value, a multiplier m_i for conclusion i is calculated

$$m_i = \pi_{j=1}^{n} (\mathrm{MIN}\ (\ 1\ ,\ (\mathrm{MAX}\ (\ 0\ ,\ (\frac{R_{ij} - I_{ij} - x_j}{2 * I_{ij}})))) \qquad (6.2)$$

The terms in this expression, R_{ij} (rejection value), I_{ij} (importance unit) and x_j (farm value in importance units), are the same as those explained in chapter 4 about the method IMAGINE. If no rejection values are given or if farm values fall outside the area, then the multiplier is equal to one.

The ultimate relevance of the conclusion is the product of m_i and R_i. A value of zero for m_i makes this relevance also zero.

In the modified version of FUZZY-DETECTOR (with rejection values), all conclusions are presented in sorted order in one overview. Figure 6.9 shows the outcome for farm F. Only the six conclusions which are most relevant as well as the conclusion with the lowest relevance ('Gross margin is very good!') are presented in this figure. The number, the name, and the degree of relevance or truth is shown for each conclusion. The position where the bar is highlighted, indicates how true or how relevant the conclusion is for farm F. The conclusion on top is certain or very relevant, while the last conclusion is never true or very irrelevant. The three conclusions that are most relevant are the same as the conclusions resulted from IMAGINE (figure 6.7). The order is slightly different however. The first two conclusions in figure 6.9 are *very* relevant or certain, while the third is relevant but not certain.

Explanation facilities for FUZZY-DETECTOR are to a certain extent comparable to those for IMAGINE. The easy readable texts and the options of graphical information about the conclusions are exactly the same. The detailed explanation is different. Figure 6.10 shows how the system comes to the conclusion 'Lower the amount of concentrates per cow; milk yield may decrease' by the method FUZZY-DETECTOR. The user which is not so familiar with the interpretation of that information can ask for help from GLOBAL-DETECTOR.

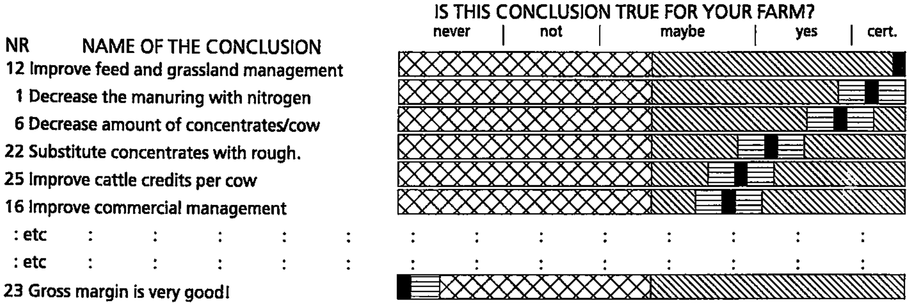


Figure 6.9 Relevance of conclusions for farm F inferred by the method FUZZY-DETECTOR. Relevance, represented by the black block, can differ from 'never' true till 'certain' true. See text for further explanation. Output from GLOBAL-DETECTOR

Figure 6.10 will only be explained shortly. One of the conclusions is presented here. The first three lines of figure 6.10 stem from figure 6.9. The conclusion is true, but not certain. Five information sources (data or variables, e.g. milk yield per cow) are needed to infer (the certainty of) the conclusion. Each variable has a condition and a weight or importance. The degree to which the farm datum (e.g. *AVERAGE* for the first variable) matches the condition from the expert (*RATHER LOW*) is highlighted on the bar. A perfect match is far right and a perfect mis-match is far left on the bar. The confluence of the positions on the bars for all five variables and their accompanying weights, result in the

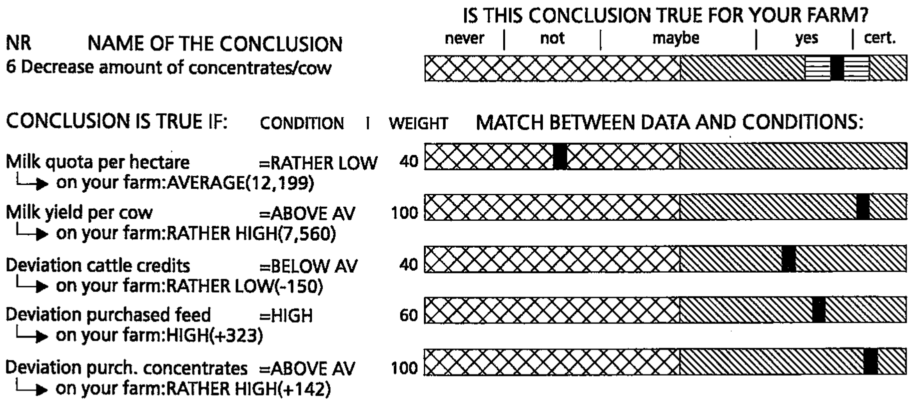


Figure 6.10 Explanation facility for the conclusion 'Decrease ...' for farm F, by the method FUZZY-DETECTOR. Output from GLOBAL-DETECTOR. (Deviation means deviation between realised and standard value)

relevance of the conclusion. The method applied by FUZZY-DETECTOR (see chapter 5) is used for the calculations.

6.4.3 IMAGINE versus FUZZY-DETECTOR

FUZZY-DETECTOR is easier to understand but less precise than IMAGINE. In this section, the outcomes of both methods are compared. The objective is to find out if FUZZY-DETECTOR is acceptable and not too rough. The extension of the method from chapter 5 with rejection values (section 6.4.2) is applied before this comparison takes place. Although not validated sufficiently (see section 6.6), IMAGINE is used as a standard to be tested against.

Data from 30 dairy farms of the accounting year 1992/93 are used in this test. These data stem from the FADN of LEI-DLO. The relevance of each of the 24 conclusions has been inferred by both IMAGINE and FUZZY-DETECTOR. To make comparison possible, the relevance g_i of conclusion i is first transformed to the interval [0,1] by $(g_i + 100)/200$.

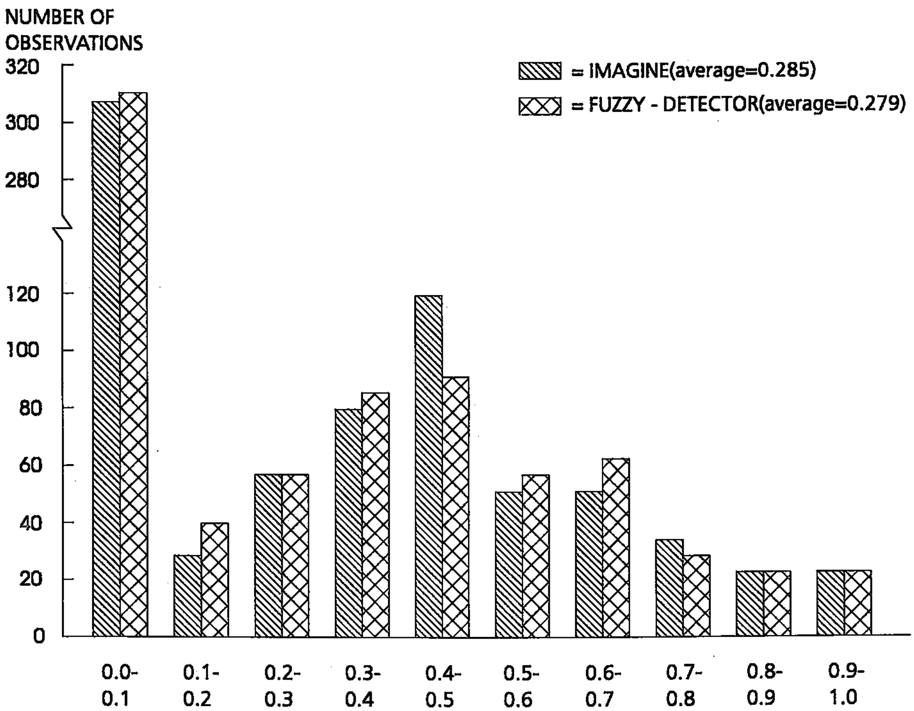


Figure 6.11 Number of observations for both IMAGINE and FUZZY-DETECTOR at different levels of relevance

There are 720 (=30*24) observations in total. Figure 6.11 gives an idea of the distribution of relevance for both IMAGINE and FUZZY-DETECTOR.

Figure 6.11 shows that nearly half of the observations have a negligible relevance (between 0 and 0.1). For the observations with a level of relevance above 0.1, the distribution between IMAGINE and FUZZY-DETECTOR seems comparable to a certain extent. With IMAGINE, many observations have a relevance between 0.4 and 0.5. The difference between the average level of relevance (0.285-0.279=0.006) can be neglected.

For each observation, the relevance between IMAGINE and FUZZY-DETECTOR is also compared. The difference D_k (k=1,..,720) between the relevance of IMAGINE I_k and FUZZY-DETECTOR F_k are obtained for the 720 observations. The absolute deviation, $|D_k|$, is calculated and shown in table 6.1.

Table 6.1 Number of observations in different classes of absolute deviations between IMAGINE and FUZZY-DETECTOR

Absolute deviation	0 - 0.1	0.1 - 0.2	0.2 - 0.3	> 0.3
Number of observations	601	96	32	1

A deviation of less than 0.1 should be regarded as equal because of the global character of GLOBAL-DETECTOR. A deviation between 0.1 and 0.2 is quite acceptable. More serious are the deviations above 0.2, which occurred 33 (=32+1) times, or in 4.6% of the observations. This is about one conclusion for an average farm. Closer examination of these conclusions showed exceptional situations, e.g. very high or very low values for one or more conditions. This is mainly caused by the fact that values of all numerical variables (conditions) are transformed to values in the interval [0,1] by the tool FUZZY-DETECTOR (see chapter 5). An example will illustrate this. Values for milk quota per hectare below 5,000 attain the value zero and above 20,000 attain the value one (=very high) by this function. So, there is no discrimination between 20,000 and 25,000 in FUZZY-DETECTOR. Both attain value one. However, this discrimination is possible in IMAGINE, because a value of 25,000 might have more influence on the relevance of a conclusion than 20,000 has.

An increase of the range [5,000 , 20,000] to [5,000 , 25,000], for example, could make the desired discrimination between 20,000 and 25,000. The transformed values are 0.75 (i.e. (20,000 - 5,000)/(25,000 - 5,000)) and 1 respectively. A drawback is that the discriminative power of all values between 5,000 and 20,000 becomes less after an increase of the range. To my opinion, the number of exceptional observations (33, or 4.6%) are acceptable enough to omit adaptation of the knowledge base of FUZZY-DETECTOR.

The exceptional observations were not caused by one or a few conclusions. The average difference between IMAGINE and FUZZY-DETECTOR for each conclusion ranged from -0.048 to 0.040. Two out of three conclusions had an absolute average difference of less than 0.02.

Neither did one or more outranging farms cause the exceptional observations. The average difference of all conclusions between IMAGINE and FUZZY-DETECTOR ranged from -0.057 to 0.049 among the thirty farms. On 27 farms the absolute average difference was less than 0.05.

Since the knowledge base for FUZZY-DETECTOR was created out of the knowledge base for IMAGINE by the knowledge engineer (without the help of the expert), it was to be assumed that the results by FUZZY-DETECTOR would differ. But from the results of comparison described above, it can be concluded that the differences are acceptable to justify the use of FUZZY-DETECTOR in the KBSs GLOBAL-DETECTOR and ENVIRONMENT-DETECTOR (chapter 7) as a good alternative for IMAGINE.

6.5 Accounting for user's need and management behaviour

King et al. (1990) stated that the low adoption rate of farm information systems is a consequence of the fact that most systems do not adequately meet the need of farmers. They blame it on the lack of understanding the managerial behaviour and information needs of the users.

On dairy farms there are not only differences in farm structure, but also manager's differences in information and decision behaviour. This is mainly caused by a distinction in the farmer's goals and his willingness to criticise and learn (Zachariasse, 1990). Farmers possess many, changeable goals.

Due to differences in information, decision behaviour and goals, each individual farmer has a specific need for information (De Hoop et al., 1988), and it is therefore a difficult task to develop management information systems that will be used on a large scale. It should be obvious that Dutch farms are too small (as contrasted with industries) to develop an information system for each of them or for a small group. So it is an impossible task to supply them with systems that match their individual goals completely and as a result, support their individual management at a very high level. To meet farmers' wishes and demands in the best way, they actually have to develop their own system. But usually skill, possibilities and time is lacking. Therefore, it is necessary to let farmers participate as much as possible (AF et al., 1992).

It should not be too surprising to see that GLOBAL-DETECTOR does not match up exactly with the individual wishes and demands. However, during the development of the system we have tried to account for users' need and management behaviour, as described below. Besides that, GLOBAL-DETECTOR is a very flexible system and the optional framework, the global character, the extensive explanation facilities and the changeability of the system makes it possible for the farmer to use it according to his needs.

6.5.1 Investigation of user's need

A survey by De Hoop et al. (1988) on management behaviour and information need at 21 Dutch dairy farms indicated the necessity for a system like GLOBAL-DETECTOR to increase the use of accounting data which is - according to Poppe (1991) - disappointing until now. Especially the findings of De Hoop et al. (1988) resulted in the formation of the requirements for (the development of) GLOBAL-DETECTOR as listed in the beginning of this chapter. Section 6.8 summarises how well GLOBAL-DETECTOR matches these requirements.

Investigation of user's need may seem advantageous, but early testing of prototypes is often preferable 1). It is difficult for farmers to express their needs for management information systems, while a prototype may be a good starting point to express criticism, to suggest recommendations for modification and to think about additional developments. A farmer must be aware of the possibilities of information technology.

6.5.2 Participation of users in the development

As an early prototype, GLOBAL-DETECTOR has been installed on the PC of some interested farmers who had little or no experience with management information systems. They used and tested it for a while and were asked to fill in a questionnaire about several aspects of the system. The answers were discussed in a meeting with these farmers, resulting in conclusions that directed further development. Most of the remarks could be implemented, e.g. need of additional graphs and tables, analysis of a couple of years with trend analysis, possibilities of simulation, system's selection of most interesting parts for analysis and the development and application of FUZZY-DETECTOR. Remarks which would lead to

---

1) In chapter 3 (section 3.3.2) advantages of prototyping are summarised.

a very detailed system were not implemented. The test has been done twice a year.

6.5.3 Flexibility of GLOBAL-DETECTOR

There are different types of decision and information behaviour among managers (Bemelmans, 1987). A flexible system like GLOBAL-DETECTOR makes it possible to present to a certain extent the needed information and decision support for these different types. Bemelmans (1987) distinguishes two different types of decision behaviour: analytical and intuitive. GLOBAL-DETECTOR has options for the calculation of different alternatives (simulation) and possesses many explanation facilities to present the analytical farmer ('thinking type') objective and rational information about for example the effect of buying quota at farm level. The intuitive farmer ('feeling type') can very quickly get data regarding his position compared to others and information on strong and weak aspects on his farm and suggestions for improvement. His own rules of thumb, intuition and experience are used to transform the data and information from GLOBAL-DETECTOR into action.

Also two different types of information behaviour are distinguished by Bemelmans (1987): perceptive and receptive. A perceptive farmer wants to be informed in general terms (main lines) and wishes to have condensed information, e.g. the strong and weak aspects. After being informed about the essentials, he might go into some details ('top-down' approach). The menu of GLOBAL-DETECTOR is structured in such a way that this is possible. A receptive farmer, on the other hand, wants all the available data (detailed) and tries to visualise the decision problem ('bottom up' approach). GLOBAL-DETECTOR can, by its flexibility, support both types of farmers.

6.5.4 User-friendliness and explanation facilities of GLOBAL-DETECTOR

The farmers who were involved in the test concluded that the system is very user-friendly, especially due to the menu structure. The fact that all necessary data were automatically read in, was also favourable. A number of farmers received GLOBAL-DETECTOR and a very short manual without any support. They were able to use the system without problem. The degree of required support seems very low.

Of special attention are the explanation facilities, which seem crucial. Good explanation makes a system clear and is therefore a major factor for future acceptability. By means of proper explanation facilities, the farmer may have an easy access to the knowledge that might be new to him. This possibility makes KBSs beneficial (Webster and Amos, 1987). The extension service may use such a system for their specialists as an aid or as an intelligent assistant, to increase the knowledge in complex areas, and to give uniform advice (Hennen, 1989). On the other hand, the knowledge and skill available from dairy extension specialists may be used for developing KBSs that evaluate dairy herd and farm management data. Significant opportunities for this approach exist in the US (Smith, 1989).

6.5.5 Different styles of farming

By its flexibility in use, GLOBAL-DETECTOR fulfills to a great extent the farmer's individual information need, decision behaviour, wishes, etc, but does not account for all aspects of the characteristics of the farmer. However, the methods and tools used in GLOBAL-DETECTOR could make this possible by interactive use and reasoning with qualitative data. In section 6.5.3, where different types of farmers were described, we have seen that different types of users may use the system in a different way.

We have been studying the possibilities to include the styles of farming in the system. Styles of farming have been defined, investigated and described by the Dutch sociologist J.D. van der Ploeg et al. (e.g. Van der Ploeg, 1993; Roep et al., 1991). Styles can be seen as ideas shared by a group of farmers regarding preferred management and development of a farm. Van der Ploeg et al. (1992) propose to give suggestions by comparing the individual farm results with the results of a group of farms with the same style. This can lead to wrong conclusions because not all aspects are analysed as the example in their publication shows. It may also lead to too limited information since the farmer does not get information about other styles (it is possible that adaptation of the style leads to a better fit of his objectives). Suggestions for improvement from the analysis of a comparison within a style may not be economic, and it may be uncertain for the farmer whether suggestions are valid for him.

The farmers who have extensively tested GLOBAL-DETECTOR wanted the system to account for the farm specific situation like stocking rate, milk quota per hectare, etc, and to use specific input-output relations as much as possible. They also wanted economic suggestions and to understand the rationale behind given suggestions, just like the way the current (flexible) version of our system does. Farmers can combine the acquired knowledge and information (from the system's explanation concerning these suggestions) with other aspects (e.g. own preferences) in a creative way to eventually come to a final conclusion about what to do. The final conclusion may therefore be different from the (original) suggestions of GLOBAL-DETECTOR, and because of this it is understandable that farmers every now and then disagree with the outcome. A computer system like GLOBAL-DETECTOR cannot give a complete list of suggestions because the system lacks data of a farmer's specific management capacity. A system has to be used in a creative way.

The concept of styles of farming is still very worthwhile and interesting, and for this thesis styles of farming have been applied for the prediction of the expected behaviour of farmers on policy options (APPROXI method, chapter 8).

6.6 Evaluation of GLOBAL-DETECTOR

Harrison (1991) uses the term evaluation to refer to all procedures applied to ensure that a model or system is appropriate for its intended use. The term is divided into verification, validation and sensitivity analysis. While verification tries to find out whether the system performs as intended after a correct implementation of its specifications, validation examines whether the intended structure is appropriate and performs with an acceptable level of accuracy (O'Keefe et al., 1988, and Harrison, 1991). Sensitivity analysis, an aspect that is not performed and that will not be discussed in this section, 'explores the extent to which outputs of a validated expert system vary when changes are made to rules in the knowledge base or to user input data' (Harrison, 1991).

6.6.1 Validation of GLOBAL-DETECTOR

For KBSs it is important to find out whether the expertise is modelled the right way. A number of methods exists for validating KBSs. The ones applied most are described and compared by Harrison (1991) and O'Keefe et al. (1988).

The expert from which the expertise was acquired, has carried out validation of the knowledge base built with IMAGINE. A difficulty concerning the validation is the unavailability of another person with a broad level of expertise which could perform such a task in this domain. Even if another expert could be found for validation, it is possible that conclusions are difficult to draw because differences in judgement might be caused by differences in expertise. A validation of the system with researchers from outside the institute has neither been carried out yet. But farmers from the test group were very content with the presented knowledge, and regarding their specific situation one may assume that they are experts themselves.

The development of GLOBAL-DETECTOR for this thesis was primarily focused on the methods used and the way how such a system should be built. The validity of the system has not been of major concern.

With data from six farms, the expert carried out a limited validation of the knowledge base built with IMAGINE. The expert was asked to indicate for each farm the strong and weak aspects and suggestions for improvement. He compared his judgement with the outcome (judgement) of GLOBAL-DETECTOR and discussed the differences with the knowledge engineer. This led to adaptations of the knowledge base. The same procedure was repeated later on some other farms.

The expert is positive about the outcome generated by IMAGINE in the KBS GLOBAL-DETECTOR. He appraised the consistency, even with very exceptional cases. Sometimes the system was more consistent and complete than himself. From his experience, we have reason to believe that the method is very acceptable to model his knowledge. An extended validation in due time would be welcome to prove the truth of this statement and throw a light on the possibilities for other domains by other experts.

In section 6.5.5 it was remarked that a farmer might disagree with the suggestions given, even if these would have been validated thoroughly enough. The reason for this disagreement is that the economic objective of GLOBAL-DETECTOR can be different from the objectives of the farmer. If the farmer is aware of this and if the outcome is explained to him (by the explanation facilities), then a non-acceptance of a suggestion must not be regarded as non-acceptance of GLOBAL-DETECTOR.

The knowledge base for FUZZY-DETECTOR has been developed by the knowledge engineer from the knowledge base for IMAGINE (6.4.2). In section 6.4.3 we have seen that results from both knowledge bases are quite comparable. There has been no test to validate the knowledge base for FUZZY-DETECTOR. The expert consulted this knowledge base regularly during the last year and he occasionally was asked to check on a questionable outcome, but he seldom suggested to make alterations.

GLOBAL-DETECTOR is a hybrid system, and the knowledge base is only one part of the system. The other (conventional) part has been validated also. The FASs values were compared regularly with the outcome of a spreadsheet. During many tests, researchers, farmers and other users were asked to pay attention to bugs, to wrong or doubtful outcome, and to wrong, missing, superfluous or unclear presentations and explanations. This led to adaptations of GLOBAL-DETECTOR.

6.6.2 Verification of GLOBAL-DETECTOR

Verification tries to find out if the system performs as intended. GLOBAL-DETECTOR has been developed in close cooperation with six farmers who tested and evaluated the early prototypes. The system was besides tested on some accountancies by accountants and by some of their clients (i.e. farmers). The system was also used and tested by some students, researchers and other farmers and organisations. The outcome of all these tests and the generally positive reactions of the users gave us the idea that the extent to which the system fits the requirements and performs as intended is satisfactory. Especially the FAS, the user-friendliness and the automatic access to the data are appreciated very much. Several users emphasised the usefulness of such a system. Some found the part which generates the conclusions (e.g. suggestions for improvement) rather difficult to understand. A substitution of IMAGINE by FUZZY-DETECTOR, which led to a better understanding, was welcomed.

A number of remarks during and as a result of the various tests, especially on presentation and explanation facilities, have implemented in the system.

A verification of GLOBAL-DETECTOR at large scale has not been performed yet. The cooperation of one or more organisations is required for this. Although several organisations are interested in the concept of GLOBAL-DETECTOR and the system itself, organisational obstacles (e.g. different data definitions and required manpower, see section 6.7.3) seem to be the main reasons that they are reserved to apply GLOBAL-DETECTOR 1). One accountancy uses FAS models at the moment and intends to cooperate with extension for giving advice based on the outcome of GLOBAL-DETECTOR.

6.7 How to make use of GLOBAL-DETECTOR

At the moment GLOBAL-DETECTOR is used on about twenty Dutch farms of the study group called European Dairy Farmers (EDF). The accounting department of LEI-DLO will use the FAS and presumably some other elements of GLOBAL-DETECTOR on about 500 dairy farms. Other accountancies have expressed their interest also, especially in conjunction with the system ENVIRONMENT-DETECTOR (chapter 7).

Identification of users should be one of the first steps of the development. GLOBAL-DETECTOR is developed in such a way that farmers can use it. Extension workers or accountants can use the system as an aid or an intelligent assistant in giving advice (suggestions). GLOBAL-DETECTOR also can generate results on paper, so that these can be added to the accounts from accountancies to increase the value of their service.

6.7.1 Use of GLOBAL-DETECTOR by farmers

Although the KBS GLOBAL-DETECTOR can be used by researchers, accountants, etc, to increase their general knowledge on dairy farm management, the system is intended to support the dairy farmers in their decision-making.

GLOBAL-DETECTOR can (by 'helicopter view') signal main striking points and give directions for improving the farm performance in future. Detailed analysis and diagnosis certainly produces better results, but by

---

1) At the moment of writing this chapter, a handful organisations showed interest to apply GLOBAL-DETECTOR (eventually in combination with ENVIRONMENT-DETECTOR).

doing so the system would grow to an unmanageable size. The farmer, with or without his advisor, can use the outcome as a starting point for further detailed analysis to detect the main causes and to take the necessary actions.

GLOBAL-DETECTOR may be used by the farmer on his own PC or by advisors (e.g. from extension service, accountancies or feed companies) on a portable PC. This will be called decentral use of the system, and the required data are supplied in files together with the programme. With decentral use it is also possible that the data are stored at the central computer of the organisation and that they are down-loaded (e.g. by mail) decentrally to the PC of the farmer for use by GLOBAL-DETECTOR. Another option is that GLOBAL-DETECTOR is only present at the central computer, and the farmer can use a terminal for the consultation. This will be called central use. Another form of central use is that GLOBAL-DETECTOR is used at the organisation by accountants or extension workers, and that the results on paper are send to the farmers. After they have received the information, they can contact the organisation for additional advice or help.

Nowadays, farmers generally have to pay for advice from e.g. extension workers. With GLOBAL-DETECTOR a farmer can save money when an advisor is consulted because both farmer and advisor are already aware of the situation, problems and possible solutions and the advisor needs less time to prepare. It is difficult to say that the role of advisors (extension) becomes less important with the introduction of a KBS like GLOBAL-DETECTOR, but their role changes and their service will be more effective and efficient. GLOBAL-DETECTOR may have perspectives to be used by extension workers as their support system.

6.7.2 Interaction with (detailed) systems

GLOBAL-DETECTOR is an aid for analysis. Due to the global character, additional advice from advisors or discussion with other farmers, e.g. in study groups, is very important to give insight in those strategies and tactics that are eligible for improvement. A further step may be the connection of GLOBAL-DETECTOR with other, more detailed, systems for analysis (e.g. Cattle credits-DETECTOR, Brée and Hennen, 1989, as well as Hennen, 1989; a system for feed and grassland management, Schakenraad et al., 1992) and planning (e.g. TACT-dairy, Zaalmink et al., 1991).

An advantage of first using GLOBAL-DETECTOR is that detailed systems can be used better later on. In an integral way (by 'helicopter view') GLOBAL-DETECTOR detects the areas eligible for further investigation. Those areas can be analysed with detailed systems. An example makes this clear. A farmer can receive the suggestion 'Improve cattle credits', which is concluded by GLOBAL-DETECTOR after examination of some characteristics of the farm based on expert knowledge. The relevance of the suggestion indicates the seriousness of the suggestion. The farmer may then call the KBS Cattle credits-DETECTOR (Brée and Hennen, 1989, and Hennen, 1989) on his PC for a detailed analysis of cattle credits. In this way the farmer uses the detailed systems more effectively and efficiently, because such systems are used when there is a problem and they can be used better since information regarding the problem is already present.

6.7.3 Transferability to other organisations

From section 6.7.1 it should be clear that organisations (e.g. accountancies) play an important role in the use of GLOBAL-DETECTOR at farm level. At the moment the system is based on automatic data input from the Farm Accountancy Data Network (FADN) of LEI-DLO. In the accounting systems of other organisations different data and calculating rules are used. This hampers the transfer of GLOBAL-DETECTOR from the LEI-DLO accounting to other ones (see chapter 2). Adaptation of the GLOBAL-DETECTOR is necessary then. The 'shell' or tool of GLOBAL-DETECTOR makes it possible to carry out drastic modifications, even in the calculation rules and the knowledge and to make a linkage with data bases of these organisations. Comparable systems in other sectors can be developed also with the 'shell'.

Some organisations, which tested the unadapted system, are positive. However, there can be some problems with transferability of data. But the organisations' disability in giving support (when a system is introduced on a farm) may be a major cause of non-acceptance. Organisations must be aware of organisational, financial and supporting problems when the system is introduced.

Quartel et al. (1992) investigated the aspects when GLOBAL-DETECTOR would be introduced at an accountancy. Some of their conclusions were:

- the KBS must fit in with the work situation of the accountancy and especially the input and the output of the KBS must fit in with existing information systems (hardware and software);
- it should be clear that the customer-related service will be better for the same price or cheaper at the same quality;
- accountants need training, especially concering the background of the analysis and calculation rules;
- the KBS must be user-friendly, this means that accountants should work easily with the KBS without much loss of time and understand what is happening.

Since a good support is an important requirement for the application of information systems (Klink, 1991), many accountancies have to extend their predominantly financial support to more technical and economic support or leave such support to other organisations (e.g. extension service). Extension service organisations can give the right support but mostly lack the automatic access to the farm data. So a good cooperation or organisational adaptation is needed.

6.8 Discussion

The requirements for the development of methods and KBSs are listed in section 1.4 and summarised at the beginning of this chapter. From our experience and after several tests we can conclude that GLOBAL-DETECTOR fulfils these requirements to a great extent. The system uses data that are already available, is aiming at tactical decision-making, gives suggestions for improvement, it gives much insight and stimulates the creativity of the farmer, takes into account the set-up of the farm, gives good standards for comparison (FAS), it fits the needs and management behaviour of many potential users to a great extent, a fast and easy maintenance is possible, it can be used by different kind of users (farmers, accountants, extension workers, etc), and the KBS and its methods and tools are applicable in other domains and with other types of data. However, this does not alter the fact that a further validation of the represented knowledge is still needed, followed by a verification of the system at large scale. Additional improvement of the system is always possible, especially in good interaction with users which can deliver very useful knowledge. In future the knowledge transfer should not be limited to one direction (i.e. from researchers and advisors to farmers), but should be bidirective. Prospects are for expert systems and specialised knowledge acquisition tools to acquire expertise from farmers to replenish the knowledge base.

GLOBAL-DETECTOR is, as the name indicates, a global system for the detection of problems on dairy farms. Farmers do not expect that information systems give the ultimate answer, but want to use these systems for support. They want to gain more insight in the aspects of decision-making, in the rationale of algorithms and in the direction of the outcome (De Hoop et al., 1988). This should stimulate the dairy farmers in their creativity. A system which is prescriptive instead of informative is therefore not desirable for the majority of farmers. The farmer must always take the ultimate decisions.

7. ENVIRONMENT-DETECTOR: KNOWLEDGE-BASED SYSTEM FOR REDUCTION OF NITROGEN SURPLUS ON DAIRY FARMS WHILE MAINTAINING INCOME

'Research on farm level shows large differences between farms in the environmental impact of their farm system and farm practice ..... Only recently these differences have become an important topic in farm management.' (Poppe, 1992)

'Models of the future will take into greater consideration the economic and environmental aspects of planned decisions. Of particular importance will be the evaluation of the possible trade-offs between these two somewhat conflicting objectives. ... Both economics and environmental concerns will demand the development and use of these systems.' (Harsh, 1990)

'De computer is eigenlijk een soort voorlichter, ... wil ik bij voorbeeld het saldo per ha gelijk houden, maar minder stikstof gebruiken, dan geeft de MILIEU-DETECTOR mij advies. Ik raak nieuwsgierig en vraag vervolgens de hulp van een deskundige. Ik weet dan precies wat ik hem moet vragen.' 1)
(Interview with dairy farmer B.Prins, (1993), 'Agrarisch Dagblad', 7(228):2)

7.1 Introduction

On many Dutch dairy farms, the input per hectare of nitrogen by mineral fertiliser and purchased feed stuffs (roughage and concentrates) increased considerably in the seventies and eighties. Relatively low feed prices versus high land prices made intensive farming practices profitable but problematic for the environment. The nitrogen output per hectare increased also on most farms, but the difference between the input and the output (i.e. the so called nitrogen surplus on the mineral account) raised. The high surpluses of nitrogen at the moment cause unacceptable high nitrate concentrations in the groundwater and surface waters and too high levels of nitrogen emission to the air (e.g. ammonia volatilisation). The Dutch government tackles the environmental problem

---

1) 'The computer is actually a kind of extension worker, ... for example, if I want to maintain the gross margin per hectare while using less nitrogen, then the ENVIRONMENT-DETECTOR gives me advice. I get curious and ask thereupon help from an expert. Then I know exactly what I must ask him (transl.WH).'

by formulating obligations and aims in policy documents and by developing instruments to reduce and control the nitrogen losses to the environment.

One of the instruments of the government that seems to be effective and acceptable is the so called 'mineral account' plus a levy system for unacceptable mineral losses. A mineral account is a statement of flows of minerals resulting in a net surplus of minerals after correction for changes in inventory (Poppe, 1992). The input or inflow of the mineral nitrogen on dairy farms stems predominantly from purchased fertilizers and feed; the output or outflow from selling livestock and milk products. The Dutch government intends to make the calculation of the mineral account compulsory for dairy farms, presumably by January 1th 1996 with a levy system when a certain acceptable mineral loss is exceeded. The administration of the input and output of minerals will lead to extra costs for all farmers, but on the other hand 'the mineral account can be a guidance for farmers to reduce the surplus of minerals' (Baltussen et al., 1993a). The dairy farmers will be informed about mineral flows, the efficiency of the mineral input and ways to improve this efficiency on their farms. Baltussen et al. (1992) advocated the mineral account since it is an instrument for economic incentive and voluntary persuasion.

Daatselaar et al. (1990) found large differences between dairy farms regarding the surplus of nitrogen per hectare. These differences were especially caused by differences in milk quota per hectare, amount of fertilisers and the level of feed and grassland management. Aarts et al. (1992) noted 'costs and benefits of measures, both in environmental and economic terms, depend strongly on specific farm conditions'. Measures to be taken by the individual farmer should therefore be farm specific, taking into account the farm's structure and performed management. In this context, Daatselaar (1989) suggested the development of a management information system, which could give farm specific advice. Such aid, eventually together with education and extension, might lead to a considerable profit for the environment without necessarily loss of income. Baltussen et al. (1992) also remarked that improvement of mineral efficiency might improve income.

The objective of this chapter is to describe the design of ENVIRONMENT-DETECTOR, a knowledge-based system (KBS) for the analysis of the nitrogen efficiency (i.e. surplus) on dairy farms and for the generation of global suggestions for a reduction of nitrogen surplus while maintaining the income as good as possible. The analysis and inferring of suggestions is based on the nitrogen account, on the farm's structure, and on the farmer's position, attitude and performed management. Farm-adjusted standards (FAS, see chapter 2) are used for positioning, and FUZZY-DETECTOR (chapter 5) is used to infer suggestions.

ENVIRONMENT-DETECTOR is described in the same chronological way as it actually works. Data from the account of farm F have been used as input for the system, and the results of the calculations are used to illustrate the KBS ENVIRONMENT-DETECTOR. Many references are made to the KBS GLOBAL-DETECTOR, because ENVIRONMENT-DETECTOR should be regarded as an extension of the system described in chapter 6.

7.2 Getting started with ENVIRONMENT-DETECTOR

The user can consult the KBS ENVIRONMENT-DETECTOR as an option of GLOBAL-DETECTOR. The technical specifications, the required hardware, the problem of uniformity of accounting data (chapter 2), the global character, and the methods used (FAS, FUZZY-DETECTOR) are the same in both systems. The reader is referred to chapter 6 for more information regarding these aspects.

ENVIRONMENT-DETECTOR requires additionally about twenty data from the mineral account, these can also be read automatically from the data base of the accounting organisation.

Once the user has started ENVIRONMENT-DETECTOR, he can choose among five options. Each will be described and illustrated in the sections below.

7.2.1 The mineral account

The first option the user can choose is the mineral account of his farm. Figure 7.1 shows this account of farm F for the accounting year 1992/93. On the left side is stated the inflow or supply of nitrogen in kilogrammes per hectare of farmland. The right side shows the outflow or removal of nitrogen. This is a statement of the flow of nitrogen, resulting in the nitrogen surplus per average hectare. On this account eight entries of inflow and four entries of outflow are distinguished 1).

The inflows of fertiliser and concentrates are generally the most important ones. Also important are the deposition (out of the air), mineralisation (additional from peaty soil compared to sand and clay) and binding of nitrogen (by leguminosae, e.g. clover). Assumptions for these are reported by Daatselaar et al. (1990) and they are taken together as the entry 'Deposition/Mineral/Binding'. This entry will not be affected by the calculations below with the arithmetical model. The entry 'Purchase fibrous roughage' is the difference between the buying and selling of roughage, corrected for changes in inventory. This may be high on farms with intensive farming practices and negative for extensive practices. The value is near zero for self-sufficient farms. The outflow

1) Organisations who want to use ENVIRONMENT-DETECTOR can change the listed entries in figure 7.1 by their preferred entries.

is predominantly the selling of milk or milk products, and to a lesser extent the balance of livestock (nitrogen in purchased minus nitrogen in sold livestock).

DETECTOR 11111/92

NITROGEN ACCOUNT PER HECTARE OF FARMLAND

INFLOW OF NITROGEN		OUTFLOW OF NITROGEN	
Purchase of fertiliser	390	Selling milk/milk product	66
Purchase of manure	0	Selling of manure	0
Deposition/Mineral/Binding	47	Balance of livestock	12
Purchase of concentrates	90	Miscellaneous	2
Purchase of fibrous roughage	19		
Purchase roughage without fibers	0	Total outflow of nitrogen	79
Purchase milk products for calves	2		
Miscellaneous	3	SURPLUS OF NITROGEN:	472
Total inflow of nitrogen	551	Total outflow and surplus	551

Press SPACEBAR......_

Figure 7.1 The mineral account for farm F. Output from ENVIRONMENT-DETECTOR

As can be seen from figure 7.1, the total inflow is 551 kg of nitrogen per hectare of farmland and the total outflow is 79 kg. The difference, 472 kg, is the surplus per average hectare on farm F. Is this surplus of 472 higher than the surplus on comparable Dutch farms? This question can be answered on the next part of ENVIRONMENT-DETECTOR.

7.2.2 The position of the farm

Farm-adjusted standards (FASs) are used to compare the farm with comparable other Dutch dairy farms (see chapter 2). FASs have been developed for the surplus of nitrogen and for the amount of fertiliser supplied. Other FASs for ENVIRONMENT-DETECTOR will be developed in due time.

Figure 7.2 shows for nitrogen surplus the position of farm F compared to average Dutch farms with the same milk quota per hectare, with the same area of feed crops (as percentage of total area), and with the same number of livestock per hectare. A distribution is drawn which indicates the number of farms at different levels of surplus. Although not fully correct, a normal distribution is assumed 1). Notice that the X-axis is turned 180 degrees and that the distance between two positions equals one standard deviation (86 kg).

The position of the farm, with a surplus of 472, is indicated by a symbol in figure 7.2. The FAS value for farm F (=376) is situated in the centre of the distribution. All farms on the right of farm F, i.e. with a surplus lower than 472, are more favourable regarding this aspect. They are denoted as 'better' farms in this figure. Seven out of eight comparable farms (88%) have a lower and more favourable surplus 2). Farms with an unfavourable surplus, or 'worse' farms, have a higher surplus and are situated left from farm F. Farmers appreciate figures like figure 7.2 to see where they are compared to their colleagues in the same situation.

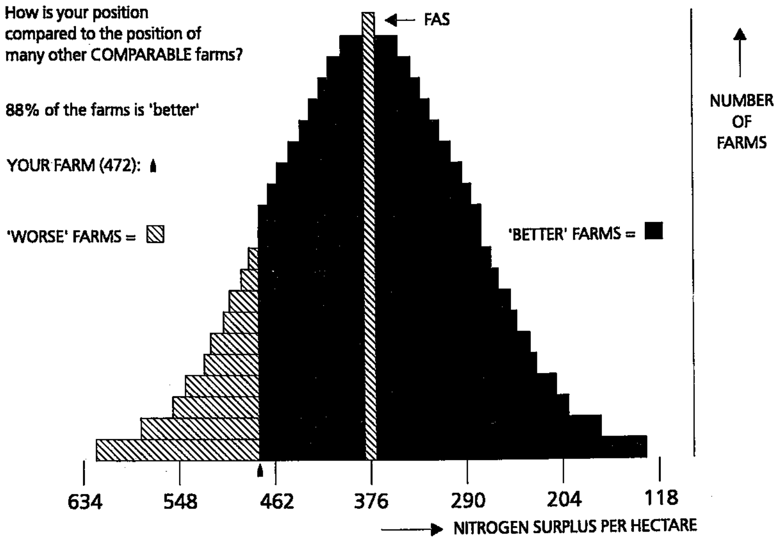


Figure 7.2 Position of nitrogen surplus per hectare farmland for farm F compared to comparable other farms. Output ENVIRONMENT-DETECTOR

1) This distribution can easily be constructed with no other information than the standard deviation of the FAS model. In due time this distribution will be replaced by the (approximation of the) real distribution, although it will require more data and programming effort.
2) Under the (incorrect) assumption of a normal distribution.

7.2.3 The position on other farms

With figure 7.2 the user knows his position with respect to comparable other farms. He may also ask for information to get insight in the relation between some variables and the surplus, or he wants to know what the average surplus is on farms with a higher milk quota per hectare but with the same values for other variables.

Figure 7.3 shows, for example, the nitrogen surplus on farms with the same area of feed crops and the same number of livestock, but with varying levels of quota per hectare. The upper line shows the FAS for the average of the highest (or 'worst') 25% of the farms corrected for the same independent variables (quota, area feed crops, number of livestock). The lowermost line represents the average of the 'best' 25% of farms. The nitrogen surplus on farm F is situated near the average of the 'worst' quarter of farms (upper line). From figure 7.3 it becomes clear that the surplus increases considerably by increasing values of milk quota per hectare. This increase is nearly 100 kg at an increase of quota with 5,000 kg milk per hectare. If the quota per hectare on farm F would have been around 20,000 kg, while the surplus and other variables remain the same, then this farm would have the position of average farms. However, from the relation in this figure it might be expected that the sur-

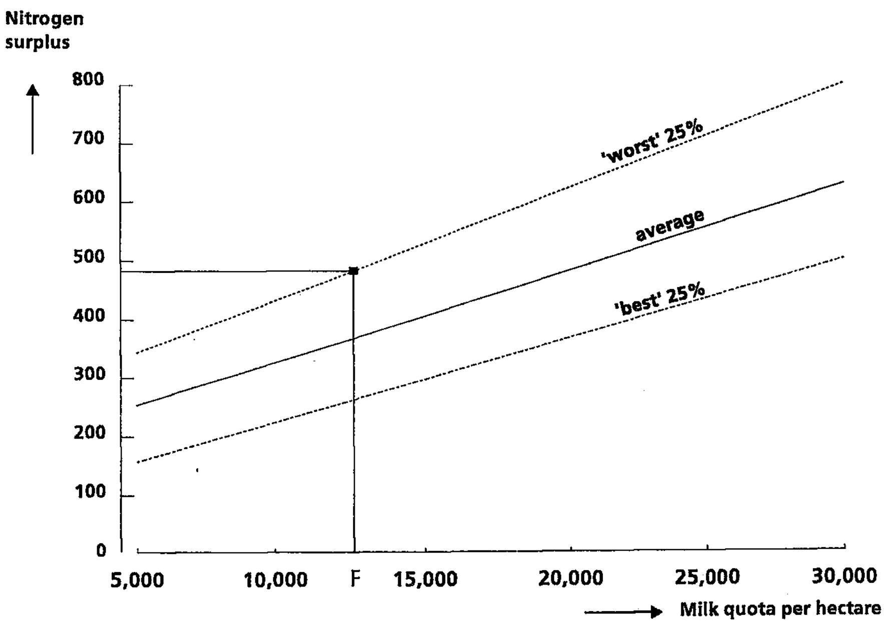


Figure 7.3 Relation between the milk quota per hectare (X-axis) and the nitrogen surplus per hectare (Y-axis) for year 1992/93; and the position of farm F. Output from ENVIRONMENT-DETECTOR

plus on farm F will increase when the quota increases. See chapter 2 for more information on how to interpret such relations from FASs.

Like figure 7.3, there are a number of other relations the farmer can ask for. They will not be described or discussed here.

7.2.4 Suggestions to decrease the nitrogen surplus while maintaining income

Up to this stage the farmer knows his position, and he might be interested to lower the surplus at the level of the average farm or even at the level of the 'best' 25% of farms. There are a number of ways to reach such a goal. The most preferable way depends on the specific situation of the farm (e.g. intensity), on the performed management and efficiency of production (e.g. does he have high costs for feeding compared to other comparable farms?), and on the effect on financial results (e.g. gross margin or income). Until now, the farmer is willing to decrease the surplus under the condition that the financial results remain the same. After 1996, when the government intends to introduce a levy on unacceptable surplusses, it is expected that many farmers decrease their surplusses (see chapter 8) because otherwise they have to pay a levy.

ENVIRONMENT-DETECTOR evaluates ten different suggestions that might decrease the nitrogen surplus while maintaining the gross margin as good as possible. The gross margin might even increase. There are also two suggestions that increase both surplus and income (e.g. 'Increase milk quota per hectare'). For the evaluation the following aspects are taken into account:

1. The farm's position with respect to nitrogen surplus compared to similar farms. A bad position will put more emphasis on suggestions that realise a great reduction of the surplus (e.g. the suggestion 'decrease amount of fertiliser').
2. The farmer's objective. The farmer can choose among three alternatives to decrease the nitrogen surplus:
   a. Hardly any reduction. Suggestions that increase gross margin are in favour;
   b. A considerable reduction. Suggestions that decrease the surplus, while maintaining or even increasing gross margin are in favour;
   c. A drastic reduction. Much emphasis is put on suggestions that realise a great reduction of surplus, while trying to maintain the gross margin as good as possible.
3. The farm's structure. An intensive farming practice, for example, with a high milk quota per hectare, large number of livestock, etc, will put more emphasis on suggestions that extensify farming.
4. The farmer's management and efficiency of production. Deviations between realised and FAS values, e.g. for cattle credits, animal costs,

quantity and cost of purchased feed stuffs, etc, are indications of good and bad management and of (in)efficient production.

5. The expected outcome when the suggestion is applied on the farm. The farm's position and the farmer's objective are used as data to create an input for the arithmetical model (see section 7.2.6). With this model the expected outcome (i.e. changes in nitrogen surplus and gross margin) can be calculated.

At the moment different styles of farming (e.g. Van der Ploeg, 1993) are not taken into account in ENVIRONMENT-DETECTOR. The methods and tools used in the system offer the possibility to include styles. But this is not done for reasons explained in chapter 6. The user can use the flexible system ENVIRONMENT-DETECTOR according to his own specific decision and information behaviour and style of farming.

Knowledge of the expert is used to develop the knowledge base of ENVIRONMENT-DETECTOR. This is done with the tool FUZZY-DETECTOR (chapter 5), which makes it possible to build a knowledge base fast and easy. Maintenance can be done with little effort.

After data regarding the five aspects above are read in by the KBS, all suggestions are evaluated by FUZZY-DETECTOR. The result of this evaluation is a presentation of all suggestions in sorted order in one overview. Figure 7.4 shows the outcome for farm F. Only the five suggestions that are most relevant are presented here. In this figure, the number, the name, and the degree of relevance or truth is shown for each suggestion. The position where the bar is highlighted indicates how true or how relevant the suggestion for farm F is. The three suggestions on top are true or might even be certainly true, while the fourth suggestion is true, and the fifth is maybe true.

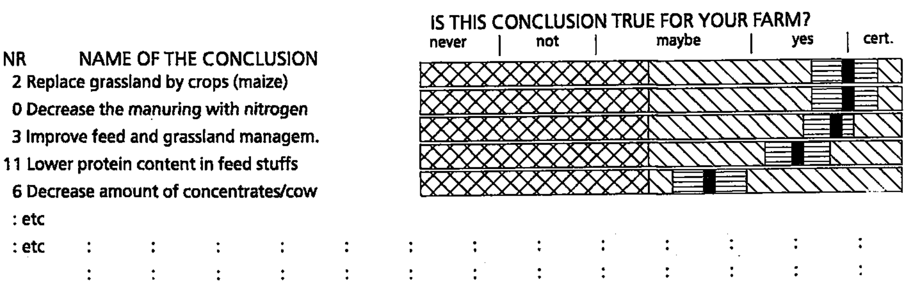


Figure 7.4 Conclusions for farm F inferred by the method FUZZY-DETECTOR. See text for explanation. Output from ENVIRONMENT-DETECTOR

Extensive information about how the expert has reached a certain suggestion can be retrieved by the user (explanation facilities of GLO-

BAL-DETECTOR). First an easy readable text is shown on screen about how the expert in general comes to such a conclusion and why this conclusion for this analysed farm has the indicated relevance. This is followed by a more detailed explanation on screen.

Figure 7.5 shows how the system in detail comes to the relevance of one of the most interesting suggestions for this farm: 'Replace [part of the] grassland by crops ([e.g.] maize)'. The tool FUZZY-DETECTOR is used for inference and presentation. A user who is not so familiar with the interpretation of that information can ask help from ENVIRONMENT-DETECTOR.

Figure 7.5 will only be explained shortly. The first three lines of figure 7.5 stem from figure 7.4. The suggestion is true or certainly true. Four information sources (data or variables, e.g. milk quota per hectare) are needed to infer (the certainty of) the suggestion. Each variable has a condition and a weight or importance. The degree to which the farm datum (e.g. *AVERAGE* for the second variable) matches the condition from the expert (*RATHER LOW*) is highlighted on the bar. A perfect match is far right and a perfect mismatch is far left on the bar. The confluence of the positions on the bars for all four variables and their accompanying weights result in the relevance of the conclusion. The method applied by FUZZY-DETECTOR (see chapter 5) is used for the calculations.

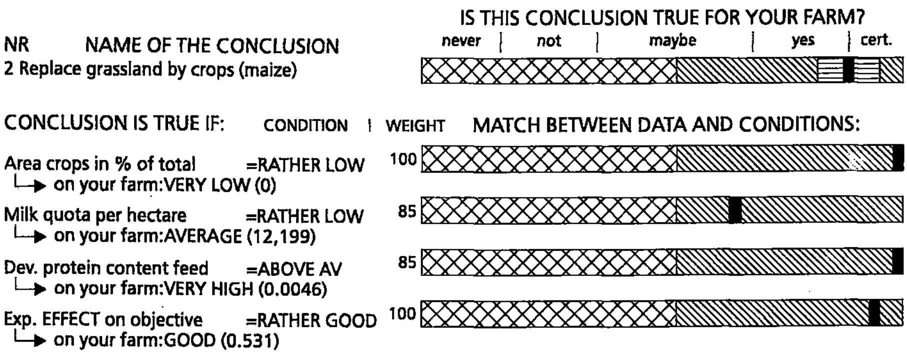


Figure 7.5 Explanation facility for the suggestion 'Replace grassland by crops (maize)' for farm F, by FUZZY-DETECTOR. Output ENVIRONMENT-DETECTOR

The last condition, 'expected effect on the objective is at least rather good' is no expert knowledge. The expert evaluates the truth of a suggestion only by the first three data. To account for the fact that the suggestion must also be interesting for the farmer, the expected outcome is included as well.

The value of the expected effect is the combination of the farmer's objective (hardly any, considerable or drastic reduction of surplus) and the expected changes in nitrogen surplus and gross margin after calculations with the arithmetical model (see section 7.2.6). The algorithm for the calculation of 'expected effect on the objective' will not be presented here.

7.2.5 Generation of tactics (packages of suggestions)

The fifth and last option of ENVIRONMENT-DETECTOR is the generation of tactics. The recommendations for the farmer, as derived in the previous section, are not separate suggestions but rather a combination of one, two or three suggestions in one package. Such package will be called tactic. The combination of the twelve different suggestions result in nearly 300 different tactics. This number will increase exponentially when new suggestions are added in future (e.g. suggestions for other environmental problems). To reduce the time-consuming calculations, only the five most relevant suggestions are combined in the tactics. This will result in 25 different tactics. Since the expert also indicated that some combinations between two suggestions are not welcome (e.g. between 'Replace grassland by crops (maize)' and 'Lower protein content in feed stuffs'), the number of tactics to be evaluated are generally less than 25. The sharp reduction in tactics is justified because the system only presents the two most preferred tactics (see below).

Each tactic, out of a maximum number of 25, is subject to the calculations with the arithmetical model (section 7.2.6). In this way interactions between suggestions are taken into account. The arithmetical model yields for each tactic the expected effect, i.e. the combination of the farmer's objective and the expected changes in nitrogen surplus and gross margin. The tactic with the highest effect is the one most preferred. For farm F, for example, this tactic is:

TACTIC1: Replace grassland by crops (maize) and
Decrease the manuring with nitrogen and
Improve feed and grassland management

ENVIRONMENT-DETECTOR selects a second interesting tactic because two tactics are to be presented to the farmer. The constraint is that the second tactic must not be too similar to the first, most preferred, tactic (i.e. TACTIC1). At the most one suggestion is allowed to appear in both tactics. We have chosen such constraint because (1) we want to present other interesting solutions, (2) the suggestions in the most preferred tactic (TACTIC1) may not be the farmer's most preferred suggestion, and (3) the farmer can always ask for a list of tactics which were also very preferable but excluded by the constraint (using 'F4:MAKE-OWN-TACTIC' in figure 7.6; see also section 7.2.7).

Accounting for this constraint, the second most preferred tactic on farm F is:

TACTIC2: Decrease the manuring with nitrogen and
Lower protein content in feed stuffs

The two most preferred tactics, TACTIC1 and TACTIC2, are presented to the farmer on a screen like figure 7.6. The contents of the two tactics, as well as the expected effects on inflow, outflow and surplus of nitrogen and the expected effects on total returns, total variable costs and gross margin are shown on the screen. These can be compared with the current values of farm F (see column 'CURRENT'). On the last three lines the differences with the current situation are shown, together with an estimated increase of fixed costs. All values in figure 7.6 are expressed per hectare.

DETECTOR 11111/92

ENVIRONMENT-DETECTOR: nitrogen account and gross margin for two tactics

		Replace grassland by crops (maize) Decrease the manuring with nitrogen Improve feed and grass-land management	Decrease the manuring with nitrogen Lower protein content in feed stuffs
	CURRENT	TACTIC1	TACTIC2
TOTAL INFLOW	551 kg N	427 kg N	456 kg N
TOTAL OUTFLOW	79 kg N	79 kg N	79 kg N
SURPLUS PER HA	472 kg N	348 kg N	377 kg N
TOTAL RETURNS	10,501 NLG	10,491 NLG	10,508 NLG
TOTAL COSTS	3,161 NLG	2,909 NLG	3,108 NLG
GROSS MARGIN/HA	7,339 NLG	7,582 NLG	7,400 NLG
EFFECT ON SURPLUS		-124 kg N	-95 kg N
EFFECF GROSS MARGIN		243 NLG	61 NLG
INCR.FIXED COSTS		148 NLG	0 NLG

F1	F2	F3	F4	F5	F10
HELP	TACTIC1	TACTIC2	MAKE-OWN-TACTIC	ORIGINAL-TACTIC	END

Figure 7.6 Presentation of the current situation and the expected effects of the preferred tactics. Output of ENVIRONMENT-DETECTOR

7.2.6 The arithmetical model

For the evaluation of suggestions by FUZZY-DETECTOR (section 7.2.4) and for the selection of the most preferred tactics (section 7.2.5), the arithmetical model of ENVIRONMENT-DETECTOR has been used. This model consists of two parts. The first part creates an input list based on default values for the second part of the model and may be regarded as pre-calculation. The second part calculates the effect on nitrogen inflow and outflow and on returns and variable costs. The arithmetical model is described briefly in this section.

7.2.6.1 Default values

The contents of the tactic (i.e. the suggestions), the farm's position with respect to nitrogen surplus, and the farmer's objective (hardly any, considerable or drastic reduction of nitrogen surplus) determine the so called default values. The way a default value is determined is explained below with an example. From this example the meaning of the default value and the role it plays in the KBS ENVIRONMENT-DETECTOR should become clear.

The suggestion 'Decrease the manuring with fertiliser', for example, can be implemented on a farm by reducing fertiliser with a certain amount (number of kgs). That amount is called the default value, or the value the KBS calculates with. When the suggestion is not given, that amount (the default value) is of course zero. When the suggestion is given, however, the KBS has to find out what amount (default value) of reduction of nitrogen fertiliser is most suitable for the farm. A function is

Table 7.1 Factors used in functions to calculate default values

	DEFAULT VALUE	FACTORS USED TO CALCULATE DEFAULT VALUE
1	Change manuring level nitrogen	Current level, nitrogen surplus, Obj. a)
2	Number hectares replaced by crops	Area of farm land, Obj. a)
3	Maize intake by cows from 1 hectare	9,000 Dutch Feed Units, no function
4	Improvement quality grass products	Current estimated production, Obj. a)
5	Improvement quantity grass products	Current estimated production, Obj. a)
6	Increase milk quota	Current quota per hectare, nitrogen surplus, Obj. a)
7	Decrease milk quota	Current quota per hectare, nitrogen surplus, Obj. a)
8	Increase concentrates/cow/year	Current level concentrates, level of concentrates on average Dutch farms, Obj. a)
9	Decrease concentrates/cow/year	Current level concentrates, level of concentrates on 25% farms with lowest level of feed stuffs, Obj. a)
10	Incr. nitrogen utilisation slurry	Current utilisation level
11	Decrease number of young stock	Current number, nitrogen surplus,Obj. a)
12	Increase area farm land (grass)	Area of farm land, Obj. a)
13	Grass production on restricted area	75% of production on non-restricted grassland, no function
14	Decrease nitrogen content feed	Current level, Obj. a)

a) Obj.= objective in reduction (1=slightly,2=moderate,3=drastic).

used to calculate the default value for a decrease of nitrogen fertiliser. This default value is a function of the current level of fertiliser, the nitrogen surplus and the farmer's objective. The default value is especially high, e.g. -250, when the current level is high, the nitrogen surplus is high and the objective is a drastic reduction. The function is based on knowledge of an expert.

The default values that belong to suggestions that are not given are set to zero. Other default values obtain a (calculated) value, like in the example above. Functions are developed to calculate default values. These functions are predominantly based on expert knowledge. Default value i can be calculated with function $f_i(F_1,F_2,..,F_n)$, where factor F_j is the value of a farm specific variable. Table 7.1 only shows the factors used. The first default value in this table, i.e. 'Change manuring level nitrogen' has been explained by the example above. Some default values are fixed, e.g. 'Maize intake by cows from one hectare' is set to 9,000. As we will see in section 7.2.7, the farmer can manually change that fixed value into his preferred one (e.g. 10,000).

The calculated default values are used for the creation of an input list for the second part of the arithmetical model (i.e. the model that calculates the expected effects, section 7.2.6.3).

7.2.6.2 Creation of the input list for the second part

The model that creates the input for the second part of the arithmetical model calculates by means of the default values the effect on milk yield per cow, on the quality of grassland products, on the intake and production of roughage, on the nitrogen content of organic manure per cow, on the nitrogen content of feed stuffs, etc.

FAS models (chapter 2) are used for the calculation of some input values. They are used for the estimation of the intake of roughage and concentrates by young stock, and for the estimation of feed production from an average hectare of farmland at changed level of nitrogen and at changed percentage of grassland.

Other algorithms are also applied. A function is used to calculate the quality of grassland products from the expected nitrogen level. Another function estimates the amount of utilised nitrogen from slurry per hectare at the new situation (i.e. stocking rate, total nitrogen level, grazing system, application of slurry).

The most important algorithms for the creation of the input list are those derived from a normative model by De Haan, and reported in De Haan (1995) and in the documentation of ENVIRONMENT-DETECTOR (Hennen and Wien, 1994). Firstly, the algorithm from De Haan (1995) for the calculation of the current genetic level for milk yield is used. The data needed for this algorithm are the current realised milk yield (fat and protein corrected), the current amount of concentrates and maize products per cow, and the current (estimated) quality of grassland products. Secondly, the algorithm from De Haan (1995) for the calculation of the (fibrous) roughage intake by cows is used. This algorithm estimates the current intake, as well as the intake in a new situation. The data needed for this algorithm are the genetic level for milk yield, the amount of concentrates and maize products per cow, and the quality of grassland products. Thirdly, the algorithm for the calculation of the expected milk yield per cow (fat and protein corrected) is used. The data needed are the genetic level for milk yield, the amount of concentrates and maize products per cow, and the quality of grassland products.

Although the reader is referred to De Haan (1995) for the exact description of these three algorithms and the rationale behind them, figure 7.7 shows the results from the third algorithm. This figure shows the effect of various levels of concentrates per cow per year on the milk yield per cow per year at two different quality levels and at two different levels of genetic milk yield. The milk yield is corrected for fat and protein. The quality level is the amount of Dutch Feed Units per kilogramme of dry matter.

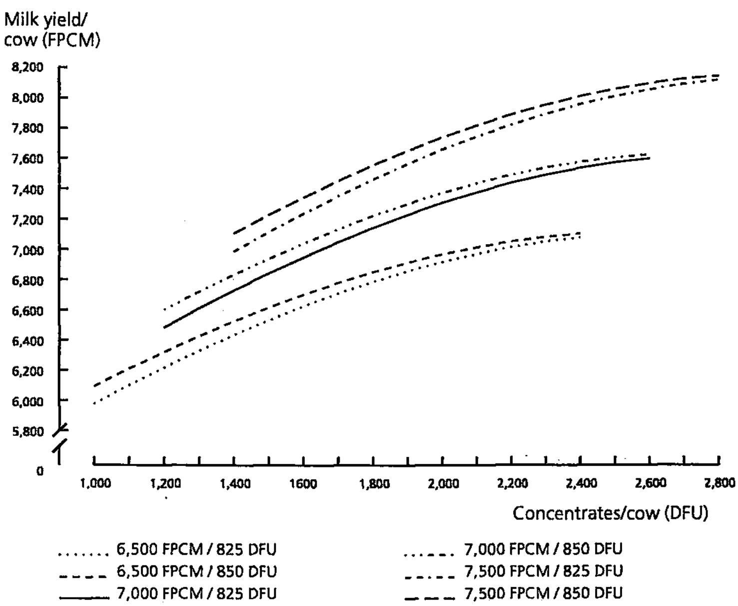


Figure 7.7 The effect of various levels of concentrates per cow per year on the milk yield per cow per year at two different quality levels and at two different levels of genetic milk yield (Source De Haan (1995), with permission)

Default and input values are illustrated in table 7.2 for the most preferred tactic from section 7.2.5 (TACTIC1, see figure 7.6) as example.

The contents of TACTIC 1 is shown on top of table 7.2. Only the default values that go along with this tactic are presented in this table. Most default values are zero and not shown here. From these default values the input values are calculated. For this tactic there is an increase in milk yield, a decrease in the purchase of roughage and an increase of nitrogen content in purchased feed stuffs. These input values are calculated with the algorithms mentioned above.

Table 7.2 Default values and input values of TACTIC1 to be used by the arithmetical model. Adapted output from ENVIRONMENT-DETECTOR

TACTIC1: Replace grassland by crops (maize) and
Decrease the manuring with nitrogen and
Improve feed and grassland management

DEFAULT VALUES THAT GO ALONG WITH TACTIC1:

Manuring level of nitrogen on grassland (incl.organic)	-100 kg N
Number of hectares replaced by crops (maize)	6 hectares
Maize intake by cows from one hectare	9,000 DFU
Improvement quality grassland products per kg dry matter	20 DFU
Improvement quantity of one hectare grassland	270 DFU

INPUT VALUES FOR CALCULATION (DERIVED FROM THE DEFAULT VALUES OF TACTIC1):

+/- manuring level of nitrogen	-100 kg N
acreage for crops	6 hectares
+/- milk yield per cow	+140 kg
+/- purchase of roughage per hectare	-515 DFU
+/- nitrogen content feed stuffs	+0.0005 kg N per DFU

DFU = Dutch Feed Unit.

When the farmer applies such a tactic, changes in the intensity of farming are to be expected on his farm. Since the quota per hectare is fixed and not affected by this tactic, an increase of milk yield per cow (from 7,560 to 7,697 kg) at the same quota per hectare will result in a decrease of the number of cows (from 96.5 to 94.8) and a decrease of the stocking rate (from 1.61 to 1.59 cows per hectare). Such changes are presented on screen.

7.2.6.3 Calculation of effects with the arithmetical model

Once all expected changes are calculated, the effect on aspects of the inflow and outflow of nitrogen and the effect on aspects of the returns and variable costs can be calculated rather easy and straightforward. The way the effects are calculated is briefly described below.

Returns per hectare

Milk receipts per hectare change with the same percentage as the change in milk quota per hectare. The milk price is assumed to be the same. Cattle credits per cow and miscellaneous returns are calculated with the FAS models for these aspects. Both are multiplied by the new stocking rate to come to values per hectare. Returns from selling of roughage are part of the variable costs.

Variable costs per hectare

Costs for concentrates per cow account for the increase or decrease of concentrates, for a change in stocking rate and for a change in price (due to different levels of maize per cow and different nitrogen contents).

Costs from the purchase of fibrous roughage or returns from selling roughage account for its increase or decrease. The price depends on the situation. If there is not enough roughage in the new situation and the farmer has to purchase roughage, then a normative price is used. If there is more roughage produced than needed by the animals, then a selling price is used (supplied by the farmer during consultation). The costs for roughage without fibers is not affected by the model, the same amount is assumed to be purchased. Costs for milk products per hectare change with the number of calves per hectare. FAS models for feed stuffs are not used directly here, but indirectly when input values were calculated (section 7.2.6.2).

Animal costs are calculated with FAS models, where the milk yield per cow is assumed to have no effect and therefore kept constant (see De Haan, 1995). Total fertiliser costs account for a change in the nitrogen level (default value), change in the nitrogen content of slurry (by changes of nitrogen level and protein content in feed stuffs), the way slurry will be applied and by a change in the percentage of grassland from total area. The costs for other minerals as well as the prices are not affected.

FASs are used to calculate miscellaneous costs. These are affected by a change in the percentage of grassland.

Changes in the amount of fixed costs per hectare are not estimated. Except an estimated increase due to a replacement of grassland by maize is calculated with FAS models. We have also accounted for compensation (income support) from the EC. ENVIRONMENT-DETECTOR gives comments on screen about the absence of other effects on fixed costs.

Inflow of nitrogen per hectare

Expected use of nitrogen fertiliser is treated the same as fertiliser costs. The kilogrammes of nitrogen in purchased concentrates and purchased or sold fibrous roughage account for increase or decrease of these feed stuffs and for changes in the nitrogen content per kilogram. The nitrogen content from milk products changes with the number of calves per hectare. All other aspects of inflow are assumed to be the same.

Outflow of nitrogen per hectare

The kilogrammes nitrogen outflow from milk products are only affected by a change in quota per hectare. The protein content in milk is assumed to be the same. The nitrogen outflow from lifestock is affected by changes in the number of young stock per cow and the stocking rate. All other aspects of outflow are assumed to be the same.

INFLOW (kg N/ha)	CURRENT	TACTIC1	OUTFLOW (kg N/ha)	CURRENT	TACTIC1
nitrogen fertiliser	390	284	selling milk/milk prod.	66	66
purchase of manure	0	0	selling of manure	0	0
deposit./mineral./bind.	47	47	balance of livestock	12	12
purchase concentrates	90	89	miscellaneous	2	2
purchase fibrous roughage	19	2			
purch.rough.without fibers	0	0	TOTAL OUTFLOW FARM	79	79
purchase milk products	2	2			
miscellaneous	3	3	SURPLUS OF NITROGEN	472	348
TOTAL INFLOW	551	427	OUTFLOW & SURPLUS	551	427

RETURNS(NLG/ha)	CURRENT	TACTIC1	VAR. COSTS (NLG/ha)	CURRENT	TACTIC1
milk receipts	9,365	9,365	concentrates	1,448	1,443
cattle credits	934	925	fibrous roughage	186	16
miscellaneous returns	202	202	roughage without fibers	0	0
			milk products	114	112
			costs for animals	863	848
			total fertiliser costs	500	366
			miscellaneous costs	50	124
TOTAL RETURNS	10,501	10,491	TOTAL VARIABLE COSTS	3,161	2,909
GROSS MARGIN	7,339	7,582	INCREASE FIXED COSTS		148

Figure 7.8 The current effects and the estimated effects of TACTIC1 on inflow, outflow and surplus of nitrogen per hectare, and on the returns, costs and gross margin per hectare. Adapted output from ENVIRONMENT-DETECTOR

The algorithms mentioned above are used by ENVIRONMENT-DETECTOR to calculate the effects of TACTIC1, our example. The results are illustrated in figure 7.8. All values in this figure, as well as the values in figure 7.9 later on (section 7.2.7), are rounded. Totals are therefore not always equal to the sum.

The major effects on the mineral account are the reductions of nitrogen fertiliser and purchased roughage. The reduction of fertiliser is not 100 kg, as might be expected from the input value (table 7.2). The value is also affected by a changed amount of manure (lower stocking rate), a decreased nitrogen content in manure, and an increase in the area of maize that requires less nitrogen. The effect on concentrates is negligible because the increase due to a higher nitrogen content in concentrates (more maize in the ration) is fully compensated by a reduction in stocking rate.

The cattle credits per hectare, on the return side, is changed by the stocking rate and by the influence of a higher milk yield per cow. The costs for concentrates remains nearly the same. The higher price, due to a higher nitrogen content, is fully compensated by a lower stocking rate. The cost for purchase of roughage and fertiliser are decreased as might be expected. The miscellaneous costs are increased, because one hectare of maize requires more seed, chemicals, etc, than one hectare of grassland. ENVIRONMENT-DETECTOR considers increase in labour, machinery, etc, as increase in fixed costs. The increase in fixed costs in our example is the extra hired labour for treatment and harvesting of maize. All other changes are caused by changes in the stocking rate.

A number of aspects are assumed not to change in the system. These are the purchase and selling of manure, the nitrogen inflow and costs of roughage without fibers, the deposition, mineralisation and binding of nitrogen and miscellaneous in- and outflow. In reality they change however.

7.2.7 Farmer's own tactic

After the farmer has indicated his objective in decreasing the surplus (hardly any, considerable or drastic), ENVIRONMENT-DETECTOR *automatically* evaluates suggestions, creates and selects tactics, creates default values and input values for the arithmetical model and calculates and presents the estimated effects. The whole procedure can be done without the presence of the farmer. In this way it is possible that accountancies can connect the system to their central data bases for the required data, run the programme for a number of dairy farms and send the results on one or two pages to the farmer.

The farmer can then ask an advisor to give additional explanations, to perform modified calculation with the chosen tactics or even create new tactics. This can be done on the (portable) computer of his advisor or on his own computer.

In section 7.2.6.1 we have seen that default values are set by the system. The farmer or his advisor can change one or more default values. When the default value for the decrease of level of nitrogen is set by the farmer to -150 instead of the value -100 from table 7.2, the system automatically creates a new input list for the calculation based on the new default value. The value of '+/- milk yield per cow' of +140 kg from table 7.2, for example, would then automatically change to a lower value (e.g. +120 kg). Such changes result in different effects on surplus and gross margin than figure 7.8 did show. The user of the system is only able to change the default values. Input values for calculation and the effects are always calculated automatically 1).

If the farmer is not satisfied with the choice of the tactics by ENVIRONMENT-DETECTOR, he can construct his own tactic out of the twelve suggestions. He has to select the option 'F4:MAKE-OWN-TACTIC' from the screen (see figure 7.6). A list of all suggestions appear on a new screen and he can choose one, two, three or even more suggestions to be included in his own tactic. Suppose the farmer chooses 'Decrease the manuring with nitrogen' and 'Decrease amount of concentrates per cow'. After his selection, a list of default values that go along with that tactic appear on his screen. He is able to change the default values set by the system. After he accepts the (changed) default values, he has no longer any influence and the system automatically infers input values for the arithmetical model and performs the calculation of the effects. Figure 7.9 shows the information the farmer gets on his screen after he has chosen his own tactic.

In figure 7.9 it is shown that the tactic the farmer has chosen reduces the nitrogen surplus by 84 kg, but decreases the gross margin as well. Although the expected outcome is less favourable than the outcome of the two predefined tactics by the system (figure 7.6), the farmer might have very good reasons to prefer his own constructed tactic.

In section 7.2.5 we have seen that ENVIRONMENT-DETECTOR automatically selects two tactics, and that the second tactic has the restriction that it should not be too similar to the first, most preferred, tactic. Because of this restriction it is very likely that there exist other tactics with better expected results than the second one. The user can ask the system for an overview of excluded tactics, also with the option 'F4:MAKE-OWN-TACTIC' from the screen (see figure 7.6). He can choose an excluded tactic, change its default values and study the results in a comparable way as above.

---

1) In earlier versions of ENVIRONMENT-DETECTOR farmers were able to change not only default values but also input values. This appeared too complex for them.

TACTIC CHOSEN BY THE FARMER: Decrease the manuring with nitrogen and Decrease amount of concentrates per cow

DEFAULT VALUES THAT GO ALONG WITH THAT TACTIC AND CAN BE CHANGED BY THE FARMER:

Manuring level of nitrogen on grassland (incl.organic)	-100 kg N
Decrease of concentrates in DFU per cow per year	-230 DFU

INPUT VALUES FOR CALCULATION DERIVED FROM DEFAULT VALUES (cannot be changed!):

+/- manuring level of nitrogen	-100 kg N
+/- milk yield per cow	-390 kg
+/- purchase of roughage per ha	+609 DFU
+/- concentrates per cow	-230 DFU

EXPECTED CHANGES ON FARM DERIVED FROM INPUT VALUES (cannot be changed!):

	CURRENT	TACTIC
Number of cows	96.5	102.0
Milk yield/cow	7,560	7,169
Stocking rate	1.61	1.70
Quota per ha	12,199	12,199

INFLOW (kg N/ha)	CURRENT	TACTIC1	OUTFLOW (kg N/ha)	CURRENT	TACTIC1
nitrogen fertiliser	390	290	selling milk/milk prod.	66	66
purchase of manure	0	0	selling of manure	0	0
deposit./mineral./bind.	47	47	balance of livestock	12	12
purchase concentrates	90	84	miscellaneous	2	2
purchase fibrous roughage	19	40			
purch.rough.without fibers	0	0	TOTAL OUTFLOW FARM	79	80
purchase milk products	2	2			
miscellaneous	3	3	SURPLUS OF NITROGEN	472	388
TOTAL INFLOW	551	468	OUTFLOW & SURPLUS	551	468

RETURNS(NLG/ha)	CURRENT	TACTIC1	VAR. COSTS (NLG/ha)	CURRENT	TACTIC1
milk receipts	9,365	9,365	concentrates	1,448	1,366
cattle credits	934	964	fibrous roughage	186	387
miscellaneous returns	202	201	roughage without fibers	0	0
			milk products	114	120
			costs for animals	863	910
			total fertiliser costs	500	377
			miscellaneous costs	50	51
TOTAL RETURNS	10,501	10,529	TOTAL VARIABLE COSTS	3,161	3,211
GROSS MARGIN	7,339	7,318	INCREASE FIXED COSTS		0

Figure 7.9 Default values (to be changed by the farmer), input values for calculation, expected changes on the farm and the current and estimated effects on inflow, outflow and surplus of nitrogen per hectare, and on the returns, costs and gross margin per hectare after the farmer's own tactic. Adapted output ENVIRONMENT-DETECTOR

DFU = Dutch Feed Units

The farmer can store each tactic and each modification of a tactic in a file. This file can always be retrieved during the same or another consultation.

7.3 Concluding remarks

In the future, farmers more and more will have to produce under the requirements of the environmental quality of products and nature and landscape. Actual information systems can be useful in giving insight in a better evaluation and fine-tuning of the input of nutrients and the output of the demanded quality of products.

Baltussen et al. (1993a) stated that the administration of the input and output of minerals will lead to extra costs for all farmers, and there will be even more costs when they will be imposed by a levy (to be expected) on the surplus of nitrogen. The farmer has to make use of the mineral account to decrease surplus and levy. But it might not be clear what kind of measures he has to take and what the expected effects on income will be. Farmers need to be supported in their decisions (see chapter 1).

Most farmers are only willing to decrease the surplus if the effect on income is not too unfavourable. The philosophy of ENVIRONMENT-DETECTOR is therefore to find ways to reduce the surplus by maintaining the income as good as possible. There are even ways to improve the income (gross margin), but this might be difficult in a futurous situation where levies and strict standards exist. Aarts et al. (1992) remarked '... more radical and expensive modifications of the farming system are necessary to meet future standards of the Dutch government for maximum allowable emissions'.

With the system measures or tactics can be found which suit the farm's position regarding the surplus, the farmer's management and objective and the farm's structure.

Knowledge from an expert in combination with the expected effects is used for the evaluation of suggestions (section 7.2.4). If only the arithmetical model (or another model, e.g. a simulation model) would have been used, the suggestions with the most favourable effects were preferred and selected. This would be a wrong procedure. Some suggestions have always favourable effects. An example is the suggestion 'Improve your feed and grassland management'. This will always decrease surplus and increase gross margin. But it is wrong to recommend this on a farm with a good feed and grassland management. On those farms there might be other, serious problems. Therefore we first have to find out (analyse) how well the farm produces compared to other farms in the same situation (FAS) and what kind of problems there are on a specific farm. For this we need knowledge modelled in a KBS as an instrument before the calculations take place.

The part of ENVIRONMENT-DETECTOR that infers the suggestions (i.e. the knowledge base of FUZZY-DETECTOR) is developed in such a way that we have tried to imitate an extension worker on the job. We assume that the extension worker comes to his conclusions in the following way. After he has studied the account, he uses his knowledge (rules of thumb) to find out which alternative suggestions (out of a large set) are most likely on this farm. As a matter of fact, he uses knowledge to match both farm data on the account and (un)favourable aspects found on the farm (compared to others, i.e. FAS) with his mental model concerning a particular alternative. The objective of the farmer, which is assumed to be known to him, is taken into account also. This is done with several suggestions. After he has found a handful of potential alternatives, he makes some calculations to have an idea how well such alternative will work out when implemented on the farm (comparable to the expected effect on objective, i.e. last condition in figure 7.5). Finally, he tells the most preferred alternatives to the farmer (comparable to figure 7.4).

ENVIRONMENT-DETECTOR serves as a so-called 'intelligent front-end' for the calculations. After the problems are detected by using expert's knowledge, the most preferred tactics are selected, and a default list and a list of input values are created. In this way it is likely that the farmer is on the right track, and from this position he can try several related options. If this is not done, the farmer will probably not know where to start and the number of combinations of input values will in fact be infinite.

The use of an 'intelligent front-end' for an arithmetical model makes ENVIRONMENT-DETECTOR a hybrid system, a system where a KBS is combined with an arithmetical model. Stone (1989) even stated that current agricultural problems cannot be solved with KBS or simulation models alone. Although this statement is too extreme, it is true that hybrid systems will be of great importance to extend the level of decision support. A system that works comparably to ENVIRONMENT-DETECTOR is HOPPER. This hybrid system described in Berry et al. (1991) uses a KBS (Consult module) to develop a list of suitable treatments and a simulation model (Economic module) to calculate the economic benefit or cost for each treatment in the list.

The calculated effects are merely approximations, although they appear as exact values in the presentation. It is not possible to make highly accurate calculations for a number of reasons:

1. The data used are readily available from the data base of the accountancy. These are generally yearly averages. It is not our intention to burden the farmer with questionaires for additional data.

Many data, that we would like to know, cannot be retrieved at all, like own feed production and feed requirements on an individual farm;

2. Farm specific input and output relations are not known, since a farm is normally not an experimental one. In ENVIRONMENT-DETECTOR we try to account as best we can for farm specific factors, but the relations are nevertheless average relations of a group of farms.

Like the KBS GLOBAL-DETECTOR, ENVIRONMENT-DETECTOR also supports decision making in a way that the dairy farmer can gain more insight in the problems and ways for improvement.

At the moment ENVIRONMENT-DETECTOR is used by one computer organisation, two accountancies and two feed factories for a test. When the experiences are positive, the organisations are intended to use the system 1). A dairy farmer can also use the system on his own computer, but it is not yet clear if this can be done with little or no support. To find this out, the system will be placed on ten to twenty computers of the Dutch members of the study group named European Dairy Farmers.

At the moment the application of ENVIRONMENT-DETECTOR is directed to the problem of nitrogen surplus on dairy farms. The methods and tools are of course not limited to this problem area. Other environmental problems on dairy farms, as well as problems in other sectors, may be tackled in a comparable way with the same methods and tools.

---

1) Test results were not known when the text of this thesis was sent away to be made up and printed.

8. METHOD TO ESTIMATE SECTOR RESPONSES BASED ON COMPOSING RESULTS FROM INDIVIDUALLY USED KNOWLEDGE-BASED SYSTEMS (APPROXI method)

"The models of the future will also make use of newer and more powerful analysis methods. For example, these models will make increased use of expert system methods." (Harsh, 1990)

"... there is not one optimal response for all the farmers to e.g. a tax on fertiliser. Each farm has its own special circumstances and input-output relations, which cause a specific optimal response."
(Baltussen et al., 1993a)

8.1 Introduction

The Dutch government tries to develop policy measures aiming to reduce a negative environmental impact of dairy farming. Before choosing one option the question is how the dairy farmers will react to the different alternative measures. LEI-DLO performs calculations regarding the economic and environmental effects of such policies on farm, regional, and national level. These studies, for example a study that estimates the effects of different levy systems on the surplus of nitrogen (Baltussen, 1992), support the government in the decision concerning what policy measure to take.

At the moment many studies in behalf of policy analysis on sector or macro level use linear programming (LP), econometric or simulation models as the approach for estimating reactions of a sector. A comparison between LP and econometric models with respect to advantages and disadvantages for the assessment of sector responses due to (environmental) policy options, has been reported in literature (e.g. Wossink, 1993; Baltussen et al., 1993a; and especially by Burrell, 1989; Bauer, 1989). Absent in both methods are the accounting for farm specific situations and the farmer's individual decision behaviour, because the whole sector is generally seen as one farm with one decision-maker. According to Baltussen et al. (1993a) "each farm has its own special circumstances and input-output relations, which cause a specific optimal response". An optimal response is not very likely in practice, because a farmer only partly knows his specific input-output relations and makes his choices not only based on maximum profit, but also or even mainly based on other factors like risk avoidance, habit, experience, etc (Elhorst and Van der Meer, 1993).

LP has been used for estimating of the effects of policies regarding the environment (e.g. Wossink, 1993; Berentsen and Giesen, 1993; Berentsen et al., 1992). With a LP approach for the whole sector there is one actor who maximises profits and there is only one input-output relation. There is no difference in goals. The base is formed by the averages of variables of (groups of) farms. Wossink (1993) developed a system based on LP models for farm categories (average of a group of farms), so that effects on policy measures can be better estimated. Such an LP model for different types of farms meets only partially the accounting for specific situations, goals and behaviour of the farmer. Mostly one input-output relation is used for the different types of farms. One LP model for each farm is not workable. Advantages of LP models are the ability to incorporate substantial changes in policy and expected technical changes.

Econometric models are also used for estimating the effects of policies regarding the environment (e.g. Becker and Guyomard, 1992; Abler and Shortle, 1992). For econometric models, elasticities and behavioural relations are estimated from empirical data (past figures). An advantage of this method is that the estimations are based on occurred adaptations caused by the behaviour, where data and developments from the past are used. The most important drawback is that since the elasticies and relations are based on rather small changes in the past, great changes in the future and new technologies cannot be taken into account (e.g. Fontein et al., 1992; Burrell, 1989). Baltussen et al. (1993a) summarises the restriction of econometric models: '...it is not easy to estimate effects of policy options with past figures of price elasticity, if these changes are very big (e.g. high taxes on fertiliser or rather big decreases in output prices) or these changes are new (tax on pollution). As a consequence of such big and/or new change there is a need for estimation of the technological change (e.g. application of manure with low emission machines, ...), and for estimation of a change in the total farm management (more efficient use of minerals, effects on the input-output relation)'. The remarks of Baltussen et al. (1993a) are partly based on the research by Burrell (1989), who discussed different methods to estimate price elasticities (regarding fertiliser) for the estimation of the effects of a sector to changes in fertiliser price. Econometric models are especially restricted when potential policy options with respect to the environmental problem are expected to result in rather big adaptations and in breaks in trends.

The choice of the method depends on the type of questions. A method for above-mentioned policy analysis, which tries to combine the strong aspects of LP and econometric models, should have the following

properties and requirements. The proposed method, and the model based on that method, should:
1. Be able to account for differences in behaviour of (groups of) farmers on various policy options, and as a result of that behaviour, be able to estimate the effects on the environment and income;
2. Estimate effects when there are big or even drastical changes, or when the changes due to the policy measure are new;
3. Account for technological change and autonomous developments;
4. Account for the farm specific input/output relations as good as possible;
5. Account for the current (structural) situation of the farm;
6. Account for general policy changes (e.g. from a result of the GATT);
7. Account for strategic aspects like structural price changes at macro level (due to changes in total production), effects because of changes in the continuity of farms, changes in the spatial distribution of farms, changes in the structure of the sector (e.g. withdrawing farmland for non-agricultural purposes), etc;
8. Use empirical data from individual farms or averages from a small group of farms. Data should be readily available (stored in data bases). The use of questionnaires for additional data should be limited. Aggregation to sector level must be possible. With the use of representative data from FADN, there is assumed that the behaviour of the sample equals the behaviour of the whole sector. The behaviour of the average farm is assumed to be different from the aggregated behaviour of individual farms;
9. Be able to combine and incorporate knowledge from various aspects in an easy way;
10. Provide insight into how behaviour and effects are derived and calculated (this insight should support policymakers in their decision);
11. Not be too big and too detailed, and therefore require little maintenance.

The method that tries to fulfil these requirements will be called the APPROXI 1) method, in accordance with the name for such a method suggested by Baltussen et al. (1993a). They mention the usefulness of a method 'that makes an approximation of reactions of a sector on alternative policy or economic options based upon the estimated reactions of individual farms' 2). In their short description of the proposal, the authors suggest to use data stemming from a representative sample of farms, e.g. from the Farm Accountancy Data Network (FADN) of LEI-DLO, and analyse these data by a KBS. They suggest to use ENVIRONMENT-

1) **APP**roximation of **R**eactions of various **O**ptions based upon farms **X**$_{\mathbf{I}}$.
2) The method APPROXI as presented in this chapter is not limited to individual farms, but may be used on data of small groups of farms.

DETECTOR, a management information system for the farmer (chapter 7), as KBS for APPROXI. In such a way maintenance, adaptations or improvements of this KBS only need to be performed once.

The proposed use of ENVIRONMENT-DETECTOR by Baltussen et al. (1993a) without adaptations is not allowed because ENVIRONMENT-DETECTOR generates suggestions for improvement (advices) and does not generate expected behaviour. A farmer does not have the same knowledge and information as the expert of the KBS. Even if each farmer would have access to the system in the near future, it is not to be expected that he would follow these suggestions indiscriminatively. To use ENVIRONMENT-DETECTOR in APPROXI for policy evaluation, the suggestions have to be transformed to expected behaviour.

The goal of this chapter is to present the APPROXI model that is based on most of the above listed requirements of the APPROXI method 1). The presented model is limited for the moment, and some requirements are only partly met. This will be discussed in section 8.4. One requirement, i.e. the accounting for strategic aspects (7th requirement), will not be fulfilled in the presentation of the model in this chapter. The measures the farmer is expected to take in the presented (provisional) model have a predominantly tactic character, i.e. measures within the current farm set-up. This is because the example of APPROXI in this chapter makes use of the KBS ENVIRONMENT-DETECTOR, and the chosen policies (section 8.2.2) make it possible to take most measures within the current farm set-up.

It is important to note that APPROXI is a method that is not restricted to the chosen KBS and policies in this chapter. The method or the philosophy can be applied in other domains with other KBSs and therefore regarding other policies. The APPROXI philosophy is especially characterised by the fact that the specific behaviour of individual farms is different from the behaviour of the average farm, that the farmer does not react in a way that maximum income is reached, and that empirical relations (comparison with FASs) in combination with knowledge are used instead of normative relations.

Our restricted first attempt in this chapter is to estimate the behaviour of farmers given the farm specific situation, attitude and policy option, followed by an estimation of the effects of the option on both the environment and income. The concept of APPROXI intends to compare effects of alternative policy options rather than making predictions.

---

1) With the presentation of the APPROXI model in this chapter we try to explain the APPROXI method. In fact, the model and method are actually the same.

It must be stressed that the model in this chapter needs further validation, which means that the estimated behaviour and effects may be different from reality. At the moment some parts of the model are validated by experts to a certain limit, and farmers are not confronted with the results yet. The other restriction on the presented model is that it does not yet account for all strategic aspects (e.g. continuity, structural changes). Because of the limited validation and absence of strategic aspects, the reader is therefore strongly urged to connect no conclusion whatsoever on the presented results of the model in this chapter. It has been our intention to explain this new method (by means of a model), illustrated with an example. The emphasis lies on the method, not on the results.

8.2 From ENVIRONMENT-DETECTOR to the APPROXI model

ENVIRONMENT-DETECTOR generates suggestions for improvement (advices), i.e. a decrease of the nitrogen surplus while maintaining income. This KBS has been described in chapter 7. APPROXI makes use of the most important parts of ENVIRONMENT-DETECTOR, i.e. the knowledge base for the generation of suggestions and the arithmetical model.

A general outline of the limited version of APPROXI is as follows. Farm data from an individual farm and knowledge stored in the knowledge base of ENVIRONMENT-DETECTOR are used to make a farm specific and economic judgement regarding the policy option. This judgement and the style of farming are thereupon used together to predict the expected behaviour of the farmer. The usage of the style of farming is typical for APPROXI. An arithmetical model is applied to calculate the effects on both environment and income. Finally, effects of individual farms are aggregated to sector level.

Processes of the general outline of APPROXI are visualised in figure 8.1. This scheme will be explained below, because it is very concise and not all relations are drawn. Following sections are devoted to explain elements of APPROXI. In section 8.1 it has already been remarked that strategic aspects are not incorporated in APPROXI. Therefore, neither in the following sections nor in figure 8.1 there will be an accounting for such aspects.

Data from an individual farm or from a small group of farms may be used. In figure 8.1 the data from an individual farm in accounting year 1992/93 are used (section 8.2.1). Some of these data are used to calculate farm-adjusted standards (FAS, see chapter 2). Objective and economic measures, or the behaviour of the farmer as a 'homo economicus', are inferred from farm data, FASs, the policy option and the knowledge stored in the knowledge base. Although not shown in this figure, the expected effect of each economic suggestion is also an information

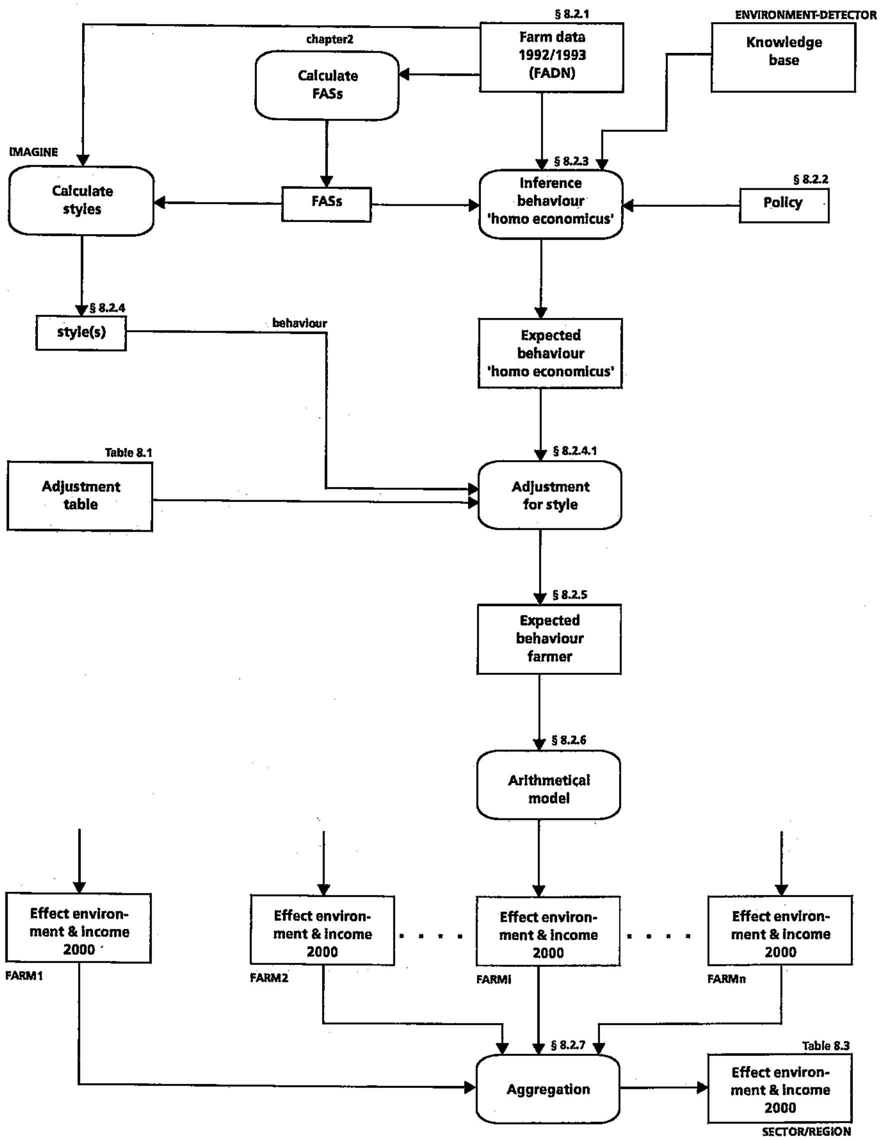


Figure 8.1 *General structure of APPROXI. The processes or sub-models are indicated by ovals. Data or information used as input for processes, as well as the results of these processes are indicated by rectangles. See text for explanation*

source to infer conclusions. The styles of farming, which are derived from farm data and FASs by using expert knowledge, shift the behaviour of the 'homo economicus' to the *expected behaviour* of the farmer. Styles of farming and the way they are calculated is explained briefly in section 8.2.4.

In this chapter the term expected behaviour is used for the list of measures derived by APPROXI that the farmer is expected to take given his farm specific circumstances and his style. These measures are predominantly measures within the current farm set-up (tactical).

Figure 8.1 shows that after the expected behaviour has been derived by APPROXI, effects on the environment and on the income (i.e. gross margin and increase fixed costs) in the year 2000 are calculated with the arithmetical model (of ENVIRONMENT-DETECTOR, see section 8.2.6). The effect on the environment is the effect on inflow, outflow and surplus of nitrogen, and additionally the effect on ammonia volatilisation and on leaching of nitrate. Section 8.2.6 describes briefly the way the effects are calculated.

Both ENVIRONMENT-DETECTOR and APPROXI show the expected effects on individual farms. Since APPROXI has to calculate responses of the sector (or region), the results from this model for individual farms have to be aggregated (section 8.2.7).

8.2.1 The data

Data can be used from individual farms or from (the average of) a small group of farms. In this chapter individual farms will be used.

APPROXI should be based on farm specific circumstances as much as possible, in accordance with some of the requirements listed in section 8.1. Therefore data will be used from existing individual farms. For this research FADN data from 300 representative specialised Dutch dairy farms from accounting year 1992/93 are used. Non-specialised dairy farms are for the moment excluded since the current version of ENVIRONMENT-DETECTOR has been only developed for specialised farms.

8.2.2 Alternative options and autonomous developments

The behaviour of farmers depends on the option the government may take. A farmer's behaviour effects his income and the environment (nitrogen surplus, ammonia volatilisation, and leaching of nitrate). Estimations for the reactions are made for the year 2000.

In this chapter two options shall be used to illustrate APPROXI:

Option 1. The Dutch government considers a policy option with a compulsory nitrogen account 1) for each farmer and a levy on a surplus above a levy-free foot. During the next years the nitrogen account can be introduced and the environmental requirements can be set stronger every year till the year 2000, the year the policy target has to be reached. As first option for APPROXI in this chapter we consider the nitrogen account compulsory and choose for a levy on nitrogen surplus of 2 NLG above a levy-free foot of 200 kilogrammes nitrogen surplus per hectare.

Option 2. The second option is absence of a levy on nitrogen surplus. In fact, this option of just 'doing nothing' by the government can be seen as the autonomous development of behaviour and effects.

In Option 2 there are only autonomous developments. Such developments also occur in Option 1, where farmers are additionally affected by the levy. Autonomous developments consist of developments as a result of policies from the past (or to be expected soon) which have to be carried out in the forthcoming years. Those policies are irreversible. In this study only a limited number of autonomous developments up to the year 2000 are assumed:

- an increase in milk production per cow per year. This increase is farm specific, and is a function of the current milk yield per cow, the milk quota per hectare and the style of farming. This function is based on the expert's expectations;
- application of low-emission techniques for reduction of ammonia emission is applied on each farm (compulsory in 2000). The obligatory covering of slurry silos is not modelled;
- slurry has to be removed from the farm when the kilogrammes of produced phosphate by livestock per hectare grassland and maizeland are beyond 110 and 70 kgs respectively (assumed costs of removal: 15 NLG per kilogramme phosphate).

A decrease of milk quota due to EC policy is not assumed and changes in prices for milk, feed stuffs, livestock, etc, are also assumed to be at the same level as 1992/93. The options will not account for strategic aspects like continuity and structural changes. This means for example that for the moment we erroneously assume that all current 300 farms still exist in 2000. In due time we will incorporate such aspects in the APPROXI model.

---

1) The nitrogen account shows the inflow, outflow and surplus of nitrogen per hectare. An example of a nitrogen account is shown in chapter 7.

8.2.3 List of economic measures

Farmers react differently to policy measures (Baltussen et al., 1993a), so the method should take into account that each individual farm has a specific structure, situation and input-output relation, that each farmer has his own goals and performed management, and should finally take into account that certain technological developments and government policies are to be expected. Knowledge (e.g. from experts, farmers) is used to estimate the expected measures these individual farmers are willing to take in the near future. This knowledge should not say what a farmer must do given his specific situation, nor what the expert himself would do, but what the farmer is expected to do given his practical circumstances, his situation, the information available to him and the capacity to utilise that information. The estimation of expected behaviour (the measures the farmer will take) is therefore not only the most important, but also the most difficult and complex aspect of the APPROXI method.

The measures the farmer will take with respect to policy options for the protection of the environment depends firstly on the economic, environmental and technical position of the farm and the resulting economic optimal adjustments, and secondly on the style of farming or the type of the farmer. The first will be determined by the knowledge base of the KBS ENVIRONMENT-DETECTOR. The second, or the style of farming, will be derived from available data of FADN.

A number of different measures are contained in the knowledge base of ENVIRONMENT-DETECTOR. These measures are:

- change the level of nitrogen;
- substitute grassland by maize;
- improve feed and grassland management;
- change the kilogrammes of milk quota per hectare;
- buy grassland with no quota;
- decrease young stock per cow;
- change the amount of concentrates per cow;
- decrease protein content of feed;
- lease restricted grassland.

From this list it is clear that most measures are tactical or within the current farm set-up. Buying farm land or buying/selling quota should be considered strategic measures since they change the farm set-up.

Farm data and FAS values are used to derive the relevance of each measure for a particular farm. For each measure these data are supplemented with information about the effects that may be expected when the measure would be implemented on the farm (chapter 7). The effects account for the levy, e.g. a price of 2 NLG per kilogramme surplus.

After the relevance of each economic measure has been inferred, a list of all economic measures with their respective relevances is set up. To come to the expected behaviour based on this list (section 8.2.5), we first adjust each relevance depending on the style of farming.

8.2.4 Styles of farming (attitude)

In chapter 6, styles of farming were explained briefly. For APPROXI, LEI-DLO has made an attempt to identify seven styles from both farm data and deviations between some data and their FAS values. For each style, e.g. for an 'economical farmer', a model was created based on combined expertise from three persons. That model contains the variables and the way to use them in order to identify the style. The IMAGINE method and tool (chapter 4) was used to model the knowledge and to calculate the relevance of each style for a particular farm. Data are used from the past year or years with respect to economic, technical and environmental performance. For APPROXI we assume that the type of behaviour in the past determines the type of behaviour in the future. APPROXI makes grateful use of the outcomes of the research on styles of farming by Van der Ploeg et al. (e.g. Van der Ploeg, 1993). The expected behaviour in APPROXI is not only determined by the style of farming (like by Van der Ploeg et al.), but especially by an economic judgement of farm specific circumstances. This makes that the behaviour in APPROXI is more dynamic than the style of farming.

The use of a knowledge model for the identification of styles is different from the data-driven identification with e.g. cluster analysis as applied by Van der Ploeg et al. A comparison between both approaches still has to be investigated. After that the eventual approach will be chosen for APPROXI.

The relevance of each style in APPROXI is a value between -100 (style not present) and +100 (style certainly present). A value of 0 means that it is not clear whether that style is present or not. We have identified the relevance of seven styles for APPROXI by IMAGINE. A style is attached to a farm when the relevance is above 20. It is possible that a farm contains more than one style, or that a style is absent (i.e a style with a relevance less than 20).

For APPROXI the seven styles that will be identified are:

1. The 'cow farmer', a farmer who gives much attentions to his cows. He does not have many cows but a high milk yield per cow. He has high costs for concentrates and high animal costs compared to colleagues (by FAS), as well as much young stock and a very high level of cattle credits. About his behaviour, which is modelled in APPROXI, we assume that he is not very willing to decrease the

amount of concentrates and the number of young stock 1). He will try to reach a very high milk yield per cow.

2. The 'machine farmer', who has more attention for the equipment than for the cows. He wants many robust cows and accepts a low milk yield per cow with low animal costs. He wants to do most of the work all by himself with his own new and modern machines. The costs for hired labour is minimal of course. This farmer is expected to be willing to grow maize and improve feed and grassland management. Compared to most others, he will not quickly increase milk quota or the amount of concentrates.
3. The 'practical farmer', who strives at a balanced farm with enough leisure time to spend with his family. The scale of the farm, expressed in the number of cows and hectares, is average. The milk yield per cow is around an economic optimal level. The costs for hired labour are quite high, since the costs for own labour and for equipment are low. Compared to other styles, this farmer is willing to improve feed and grassland management but it is assumed unlikely that he will grow maize. If the protein content in feed stuffs is high, then he will try to reduce that content.
4. The 'economical farmer', who saves costs and has a high solvability. His farm has a low stocking rate with a low producing herd, has a low level of nitrogen on grassland, low amount of concentrates and low animal costs. To save costs for hired labour, he works many hours (high calculated costs for own labour). This farmer is willing to decrease the level of nitrogen and concentrates when these are high. It is not very likely that he will increase the area of farmland or the milk quota per hectare.
5. The 'grassland farmer', who wants to have maximal production of grassland products with good quality. He has high costs for equipment and hired labour, high level of nitrogen fertiliser, and high costs for chemicals and seed for sowing. Compared to other farmers (by FAS) with the same intensity, etc, his type has low costs for purchasing of roughage and concentrates. This farmer tries to reach a high grassland production in the future, which means that it is not very likely that he will decrease the level of nitrogen much. Instead he will improve feed and grassland management even further. The amount of concentrates is not to be expected to increase much.
6. The 'environment farmer', a farmer who is environment-minded and uses low levels of nitrogen, phosphate and potassium fertiliser. The stocking rate, amount of concentrates, the protein content in purchased feed stuffs and the number of young stock are all low on his farm. The result is a low nitrogen surplus per hectare. The behav-

---

1) Knowledge with respect to the remarks about the expected behaviour in this section is supplied by J.J.F. Wien from LEI-DLO and modelled in APPROXI. See also table 8.1.

iour of this farmer is oriented to low emissions of nitrogen: decrease nitrogen and concentrates even further, improve feed and grassland management, eventually increase the area of farmland, and growing of maize. He will only increase milk quota if his farm is very extensive.

7. The 'fanatic farmer', is a farmer who has the opinion that the farm must be large to survive. He had spent much money to enlarge his farm by buying quota and land. His solvability is therefore rather low. This farmer does not quickly decrease the amount of concentrates or the number of young stock. He has a positive attitude to growing maize.

8.2.4.1 Adjustment of the measures

The style of the farm is used to adjust the relevances of the measures (section 8.2.3), which have been derived by the KBS ENVIRONMENT-DETECTOR, in a deterministic way. The objective of this adjustment is to account for the style of farming because the final behaviour (section 8.2.5) is not only rational and economic but also influenced by the attitude of the farmer (i.e. style). Table 8.1 shows the adjustment values. These values in table 8.1 are based on knowledge 1), and they are not yet validated. In this table adjustment values are only shown for two measures. Adjustment values are specific for each style and each measure.

The adjustment procedure is very easy. An adjustment value from table 8.1 for the concerning measure and style is added to the relevance of the measure. The result is bounded to the interval [0,1]. An example will be used to explain these adjustments. Suppose the style of our example is 'grassland farmer', and the relevance of the measure 'Decrease level of nitrogen' (as a result of calculations with the KBS ENVIRONMENT-DETECTOR, section 8.2.3) is 0.75 (i.e. rather relevant). This value is added to the adjustment value from table 8.1 for this style and suggestion: -0.5. The measure adjusted for attitude or style is now 0.75 + -0.5 = 0.25, which might be interpreted as a rather unlikely measure to be taken by the farmer. If the style would have been an 'economical farmer', then the measure adjusted for style would have been 0.75 + 0.5 = 1.25. Bounded by the interval [0,1] will give 1 as the relevance, meaning that it will be certain that the farmer will take that measure.

The adjustment values from table 8.1 are also used for the calculation of effects on individual farms with the arithmetical model (section 8.2.6 and figure 8.2).

1) The adjustment values in this table are supplied by J.J.F. Wien from LEI-DLO. His knowledge is partly based on the research on farm styles by van der Ploeg et al. from Wageningen Agricultural University.

Table 8.1 Adjustment values for the transformation of the relevances of the measures (derived by the KBS ENVIRONMENT-DETECTOR) to the relevances when accounted for attitude or style (positive values increase and negative values decrease the certainty factors). See text for explanation

Measures	STYLE OF THE FARM							
	cow	machine	practical	economical	grassland	environment	fanatic	not known
Decrease level of nitrogen	0.1	-0.2	0.2	0.4	-0.5	0.5	-0.2	0
Substitute grassland by maize	0.2	0.3	-0.5	-0.2	-0.2	0.3	0.4	0
etc								

In the APPROXI model it is possible that a farm may possess more than one style. The eventual style of a farm is a convex combination of these styles. We account for all styles which are relevant on a farm, and not just the style with the highest value for style-relevance.

When more than one style has been identified (i.e. relevance > 20), the adjustment values are calculated by the weighted summation in the following way:

$$\text{ADJUST-C}_i = \frac{\Sigma\ \text{ADJUST}_{ij}*\text{STYLE\_REL}_j}{\Sigma\ \text{STYLE\_REL}_j} \tag{8.1}$$

where ADJUST-C_i = adjustment value of combined styles for suggestion i

ADJUST_{ij} = adjustment value for suggestion i and style j (table 8.1)

STYLE_REL_j = relevance of style j on farm F; this relevance is > 20.

We have not used all styles, but only styles with a relevance of more than 20 because this makes it easier to define the eventual combined style of the farm. With expression (8.1) it is assumed that a style with relevance 100 is twice as important than a style with a relevance of 50. The truth of this assumption needs to be investigated by validation.

Expression (8.1) will be explained with an example. Suppose two styles are identified for farm F: 'economical farmer' with relevance 85 (STYLE_REL_4=85), and 'grassland farmer' with relevance 40 (STYLE_REL_5=40). In table 8.1 the adjustment values for these two styles can be found: the adjustment value of the 'economical farmer' is 0.4 for the suggestion 'Decrease level of nitrogen' ($ADJUST_{1,4}$=0.4, first row in table 8.1), and the adjustment value of the 'grassland farmer' is -0.5 for the same suggestion ($ADJUST_{1,5}$=-0.5). With expression 8.1 it can now be calculated that the combined adjustment value ADJUST-C_1 for 'Decrease level of nitrogen' is ($ADJUST_{1,4}$*$STYLE\_REL_4$ + $ADJUST_{1,5}$*$STYLE\_REL_5$) / ($STYLE\_REL_4$*$STYLE\_REL_5$) = (0.4*85 + -0.5*40)/(85+40) = 0.11. If the relevance of the suggestion 'Decrease level of nitrogen' would have been e.g. 0.67, then the measure adjusted for style is 0.67+ADJUST-C_1 = 0.67+0.11 = 0.88.

So we have proposed a model to adjust suggestions (measures) derived by a KBS (i.e. ENVIRONMENT-DETECTOR) by using information of the style(s) of the farm. The adjusted result is the likelihood of taking that measure by the farmer (farmer's behaviour). This way is hypothetical, further validation has to prove its correctness.

8.2.5 Formation of the expected behaviour

After adjusting for the style of farming, or the attitude of the farmer, each measure has a new relevance for the farm. It is now assumed that each (adjusted) measure with a relevance above 0.5 is applied by the farmer. These measures together will be called the expected behaviour. If the behaviour contains no measure, the farmer will only follow autonomous developments.

The expected behaviour is assumed to take place just after the policy option becomes effective, and is fully implemented by the year 2000.

The APPROXI model as described in this chapter assumes that the farmer will execute the expected behaviour. However, it might be possible that a farmer is unable to take a measure even if he wants to. There is assumed that one measure, namely the lease of grassland with restrictions ('birdland'), might not be readily available in the neighbourhood of the farm. We assume that one out of three farms has the opportunity to lease. This is not modelled stochastically, but in a deterministic way. The number of the order in which the farm entries the model, is divided by three. If the result of the division is an integer, then the likelihood (i.e. relevance) remains the same, otherwise it will be zero and will be excluded from the expected behaviour.

8.2.6 The arithmetical model: calculation of effects on individual farms

The expected behaviour will be subject to an arithmetical model for the calculation of the effects. This model is quite comparable to the arithmetical model of ENVIRONMENT-DETECTOR (see chapter 7), only a few extensions are made. These extensions are the incorporation of autonomous developments, the accounting for the style of farming, and more effects are calculated and shown (e.g. ammonia volatilisation).

An example will be used to illustrate how the style of farming influences the effects. Figure 8.2 shows with how much kilogrammes different farmers are expected to decrease nitrogen. The decrease is larger on farms with a higher current nitrogen surplus or on farms that use more nitrogen fertiliser. It is also shown that in the same situation an 'environment farmer' decreases the level more than a 'grassland farmer', due to a different attitude between these styles. Adjustment values from table 8.1 are used for this.

Figure 8.2 represents expert knowledge, and can easily be implemented in APPROXI with some simple algorithms. The function for the expected decrease of nitrogen level shown in this figure is

expected decrease = f(current N surplus, current N level, style)

Similar functions for other pieces of knowledge are modelled in the same way.

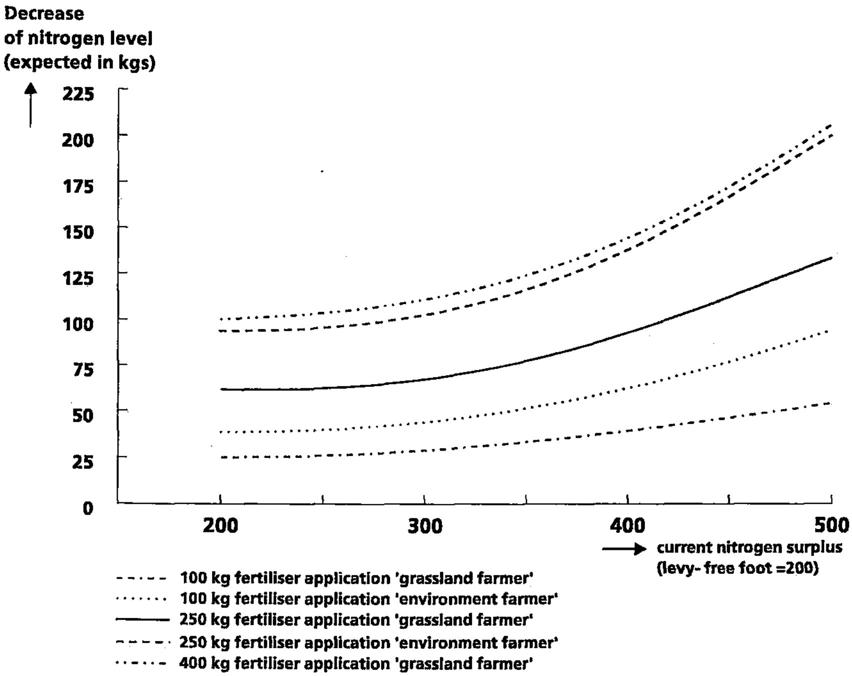


Figure 8.2 Expected kilogrammes decrease of nitrogen level at different current levels of nitrogen surplus (with a levy-free foot of 200), at different current levels of kilogrammes fertiliser application per hectare, and for two different styles. Based on knowledge from an expert

The output of APPROXI is the expected effect as a result of the expected behaviour on a particular farm. The output consists of the following elements for each individual farm:

- the style(s) and the relevance of the style(s);
- the expected behaviour (i.e. all measures with a relevance above 0.5);
- the data that have been used as input data for the arithmetical model, e.g. an increase of milk quota of 20,000 as a result of behaviour;
- the current values, the expected values from autonomous development, and the expected values from autonomous development in combination with the policy behaviour for:
    * milk quota per hectare and area of land
    * kilogrammes of nitrogen fertiliser
    * kilogrammes nitrogen surplus
    * kilogrammes nitrogen from ammonia volatilisation
    * kilogrammes nitrogen from leaching of nitrate
    * gross margin per hectare (levy included)
    * amount of yearly additional costs (e.g. quota costs)
    * costs for removal of slurry when the kilogrammes of produced phosphate on grassland and maizeland are beyond 110 and 70 respectively (15 NLG per kilogramme);
- the weighing-factor from the FADN of LEI-DLO (see section 8.2.7).

The calculation of ammonia volatilisation and leaching of nitrate is done by regression models estimated from the "Stofstromenmodel" (nutrient flow model; Van der Veen and Dijk, 1993), a model currently used by LEI-DLO and AB-DLO. These regression models, which are used in APPROXI, are not validated.

8.2.7 Aggregation of results

The representative farms of the FADN, which we use in APPROXI, are a sample from all specialised Dutch dairy farms. Thus each farm from the sample represents a number of Dutch farms. This number is the weighing-factor to aggregate to the sector level.

Aggregation is done by weighted summation in the following way:

$$\text{EFFECT}_i = \frac{\Sigma\ \text{INDIVIDUAL-EFFECT}_{ij} * \text{WEIGHING-FACTOR}_j}{\Sigma\ \text{WEIGHING-FACTOR}_j} \qquad (8.2)$$

where EFFECT_i = the aggregated effect i, or the sectoral effect
$\text{INDIVIDUAL-EFFECT}_{ij}$ = the individual effect i on farm j
WEIGHING-FACTOR_j = the weighing-factor of farm j

8.3 Results from the APPROXI model

The APPROXI model as described in section 8.2 is an extension of ENVIRONMENT-DETECTOR and has been programmed on a PC in computer language muLISP. The data that have been described in section 8.2.1 (i.e. 300 farms from FADN for year 1992/93) were used. The model has been run for all individual farms under Option 1 (i.e. the levy option, section 8.2.2), followed by a run for all individual farms under Option 2 (i.e. the autonomous development). For each option the effects were aggregated by APPROXI and placed in a comparison table as shown in table 8.2 and 8.3. The results as presented in these two tables are provi-

Table 8.2 The number of farms from the sample (300 representative and specialised dairy farms from FADN) that take measures, the expected number of farms in the Netherlands that take these measures, and the average contents of each separate measure for Option 1 and Option 2. (Adaptations are estimated for the year 2000, compared with the base year 1992/93). Output from the APPROXI model

Measures	OPTION 1 (levy on nitrogen)			OPTION 2 (autonomous)		
	number of farms		av.adaption of group	number of farms		av.adaption of group
	sample	Netherl.		sample	Netherl.	
Decrease level of nitrogen	127	8,844	-133 kgs N fertiliser	62	4,284	-78 kgs N fertiliser
Increase level of nitrogen	20	1,544	10 kgs N fertiliser	61	4,231	55 kgs N fertiliser
Change grassland with maize	183	12,120	4.3 hectares	169	11,270	4.2 hectares
Improve feed and grassland managem.	194	12,988	28 DFU quality increase 837 DFU quantity increase	183	12,173	28 DFU quality increase 839 DFU quantity increase
Lower milk quota per hectare	8	502	-26,375 kgs per farm	8	502	-21,500 kgs per farm
Increase milk quota per hectare	115	7,322	47,052 kgs per farm	134	8,566	61,910 kgs per farm
Buy grassland with no milk quota	30	2,096	4.6 hectares	29	2,017	4.7 hectares
Decrease number of young stock per cow	34	2,224	-0.15 livestock units	26	1,788	-0.06 livestock units
Decrease protein content of feed	176	11,621	-0.005 grammes N/kg DM	170	11,203	-0.002 grammes N/kg DM
Decrease amount of concentrates	122	8,360	-386 kgs per cow per year	120	8,189	-388 kgs per cow per year
Increase amount of concentrates	12	923	264 kgs per cow per year	12	923	264 kgs per cow per year
Lease restricted land	51	3,006	4.2 hectares	42	2,488	4.1 hectares

sional and merely indicative due to a not yet completed validation, absence of verification, and absence of strategic aspects (see section 8.1). The version of APPROXI that produced these results is based on knowledge, supplemented by relations developed at LEI-DLO and findings from the literature.

Table 8.2 presents the number of farms from the sample of 300 farms that are expected to take the various measures for both Option 1 and Option 2. For example, if Option 1 (i.e. levy) will be brought into effect by the government, then it will be expected from the APPROXI model that 127 farms out of 300 decrease the level of nitrogen with on average 133 kgs by the year 2000. Multiplied by their weighing-factors, these 127 farms represent 8,844 specialised dairy farms in the Netherlands. It is not surprising that under Option 2 (i.e. autonomous development) this number of farms, as well as the expected adaptation (i.e. average content), is much less. The 62 farms that decrease the level of nitrogen by 78 kgs will predominantly do that for economic reasons. The style of farming affects also both numbers.

For all measures the value of the content depends on the style of farming. For some measures the value of the content depends also on the difference between the surplus and the levy-free foot (e.g. regarding 'Decrease level of nitrogen') or on the characteristics of the policy option (e.g. regarding 'Decrease number of young stock per cow').

For most measures the number of farmers that are expected to take them, as well as the average content or adaptation, are different for both options. However, no differences between the options are found for the measures 'Lower milk quota' and 'Increase amount of concentrates' according to the (provisional) APPROXI model. The measures 'Increase milk quota per hectare' and 'Increase level of nitrogen' may seem conflicting with the policy option that tries to reach a reduction of nitro-

Table 8.3 Differences in effects between Option 1 (a levy) and Option 2 (autonomous developments) for some technical, environmental and economic aspects. Output from APPROXI

EFFECTS	OPTION 1 IN 2000 COMPARED TO OPTION 2 IN 2000
Milk yield	35 kgs per cow LOWER
Milk quota	9,747 kgs per farm LOWER
Area of farm land	0.1 hectares LARGER
Kgs nitrogen fertiliser	38 kgs per hectare LOWER
Kgs nitrogen surplus	40 kgs per hectare LOWER
Kgs ammonia volatilisation	10 kgs per hectare LOWER
Kgs leaching nitrate	22 kgs per hectare LOWER
Gross margin	8,757 NLG per farm LOWER
Net profit	3,863 NLG per farm LOWER

gen surplus. Since the main goal of farming remains a high income, some farmers will apply such aspects. Probably in combination with measures that decrease surplus. It is our intension to incorporate in future more of such income increasing measures in the APPROXI model.

Table 8.3 presents the differences in effects between the two options. The difference in milk yield per cow is explained by different behaviour regarding the application of nitrogen fertiliser (table 8.2). The difference with respect to milk quota per farm and area of farm land is explained by the reaction in table 8.2. The environment is expected to be better off with a levy on surplus (Option 1), as might be expected. The nitrogen surplus, ammonia volatilisation and leaching of nitrate per hectare will be lower. The expected average for nitrogen surplus (245 kgs, not shown) for Option 1 is higher than the value of the levy-free foot (200 kgs) for this option. Although a number of farmers have a surplus below the foot, and do not have to pay the levy, most farmers take the money to pay for the levy for granted. Maybe the saving of money for the levy by further reducing the surplus counts for little compared to managerial changes, increase in risk, decrease of production and income or a change in attitude.

The decrease in income due to the levy is expected to be 3,863 NLG in the year 2000, compared to the situation in 2000 when the levy is not affected (Option 2 or autonomous development). Part of that decrease is explained by the levy paid, but the other part (about 1,200 NLG) is

Table 8.4 The farm characteristics, the expected behaviour under Option 1, and the expected effects (by APPROXI) for three comparable farms

CURRENT SITUATION:	FARM 1	FARM 2	FARM 3
Milk quota per hectare	10,078	9,448	9,602
Area of farm land	22 ha.	22 ha.	17 ha.
Milk yield per cow	6,897	6,670	6,760
Number of young stock	0.32	0.34	0.33
Stocking rate	2.10	2.10	2.08
Style of farming	economical	none	environment
Expected behaviour (i.e. measures that farmers take)	nitrogen:- grow maize feed&grassland:+ young stock:- protein:-	grow maize feed&grassland:+ protein:- quota:+	nitrogen:- feed&grassland:+ concentrates:- 'birdland'
EXPECTED EFFECTS:			
Decrease nitrogen fertiliser	108	36	60
Decrease nitrogen surplus	166	104	110

because the farmers took measures and changed their management because of the restriction.

There are differences between different individual farms under the same option. The structure of the farm, the input of feed and fertiliser at the moment, the surplus of nitrogen at the moment, the performed management of the farmer and the style of farming all cause a diversity of reactions (or measure they take). Even if the farm structure is the same, farmers may react differently due to differences in style of farming and management. Table 8.4 shows the characteristics of three comparable farms, the expected measures when a levy will be applied (Option 1), and some expected effects.

Although the structure of the farms in table 8.4 are quite comparable, the expected behaviour is not the same. The measures farmers take depend on the style of farming, on the current levels of nitrogen surplus, nitrogen fertiliser and concentrates, on the grassland production, etc, and on the management (deviation from FAS). Different measures also cause different effects, as shown in table 8.4.

8.4 Discussion

This chapter has shown that there are possibilities to use a KBS - which was initially intended to support the management of the dairy farmer - for the estimation of sector responses on policy options according to the APPROXI method.

In the introduction of this chapter (section 8.1) a list of requirements were presented for APPROXI. These requirements represent the philosophy of APPROXI. However, the application presented in this chapter has not met all the requirements. The possibility of the development of APPROXI according to all requirements has to be proven.

In this section the degree to which the APPROXI model in this chapter fulfils the requirements from section 8.1 are discussed. The same list of requirements is listed below.

1. The APPROXI model is able to estimate the individual and predominantly tactical (i.e. within the current farm set-up) behaviour of the farmer on the policy options of the example in this chapter. As a result of that behaviour APPROXI can estimate the effects on the environment and income. Other options that are comparable to the levy option used in this chapter, e.g. a levy-free foot of 100 kgs or a levy-free foot that depends on farm characteristics, can also be handled by the model.
2. New changes like the levy on nitrogen surplus can be modelled with the APPROXI method. Although not proven in this chapter, the methods are likely to be flexible enough to model also big and drastical changes.

3. Technological change (e.g. application of low-emission techniques, genetic improvement of milk yield) and autonomous developments can be modelled by APPROXI. One of the options we used as example is an autonomous development (Option 2).
4. Input/output relations are not known for individual farms. But by the use of FASs, these relations are made as farm specific as possible (chapter 2).
5. APPROXI accounts for the current (structural) situation of the farm. Behaviour as well as effects depend on the situation of the individual farm.
6. In the presented model it is not proven that APPROXI can account for general policy changes (e.g. from a result of the GATT), but changes in prices or quota can be incorporated easily in the model. APPROXI estimates the behaviour and calculates the effect in one step from e.g. 1992 to 2000. Since there is no iterative process, gradually changes cannot be taken into account for the moment. However, such changes can be modelled when strategic aspects (next requirement) are taken into account.
7. The current version APPROXI, i.e. the version presented in this chapter, cannot account for all strategic aspects like structural price changes, effects because of changes in the continuity of farms, etc. It is our intention to extend the current model for this requirement.
8. Representative data from FADN are used so that we were able to estimate sector responses from individual responses. Data were readily available from the data base and no additional data were required.
9. We are of the opinion that the tools IMAGINE (chapter 4) and FUZZY-DETECTOR (chapter 5) make it possible to combine and incorporate knowledge from various aspects in a fast and easy way.
10. These tools make it also possible to provide insight into how behaviour and effects are derived and calculated. However, this insight can only be gained for individual farms.
11. The maintenance of the model is rather low. The model is not too big and too detailed, and the knowledge base developed with either IMAGINE or FUZZY-DETECTOR requires low maintenance.

From the evaluation of the requirements, it can be concluded that not all requirements of the APPROXI philosophy are fulfilled with the example in this chapter. But the model is satisfactory enough for the policy options chosen, except that the necessary strategic aspects (7th requirement) are not yet incorporated. The measures the farmer is expected to take in this model are predominantly tactic (i.e. within the current farm set-up).

Although the presented model in this chapter fits the requirements quite well, there is still further validation required. Especially the adjustment values (table 8.1), which influence the behaviour of the farmer according to his style, need special attention. The validation of the behaviour of the farmer and of the APPROXI model is necessary because:

- there is a general lack of knowledge and understanding of the farmer's decision-making and behaviour. Gaining more insight in this is not only important for the validation and adaptation of the APPROXI model, but also for other research topics;
- the knowledge of the expert may be inadequate, especially concerning the farmer's reaction on new or drastic measures. The knowledge used in the system must indicate how the farmer in a situation would react, and not how the expert himself would react in the same situation.

Data from the past, and the 'behaviour' found in those data, can be used to validate parts of APPROXI. In this way calculated results by APPROXI can be compared with the behaviour and effects that really happened. However, the modelled knowledge and relations in the current prototype are strongly linked with the chosen levy option as policy measure. Since such option is new and since drastic changes are to be expected, a validation of this restricted model with past figures is limited. Additional knowledge and relations have to be implemented in the current model for other (types of) policy measures or for use in other domains.

Validation can also be done with questionnaires or interviews. A side-effect of such validation is that we may gain more knowledge and understanding of farmers' decision-making and behaviour. Although validation with questionnaires and interviews are useful, some limitations have to be taken into account:

- some farmers might not know at the moment how to react on a policy measure in future, especially when a measure is new and drastic. Even though a farmer could indicate a certain reaction, there still might be a chance that he dissents;
- circumstances may influence and change farmer's behaviour, e.g. the introduction of other policy measures or the introduction of a management support system like ENVIRONMENT-DETECTOR on the farm;
- changing attitudes of the society might interfere with the behaviour;
- it may be difficult to find out how a certain behaviour turns out. Suppose for example that a farmer indicates he will decrease the level of nitrogen. For the model it must also be known with how much kilogrammes he will do so and what the effects will be on grassland production, milk yield, etc. This depends amongst others on the unknown capacity of the farmer.

Policy makers might have trouble accepting the outcome since a KBS (or expert system) is not yet an accepted phenomenon. The validation is therefore important. Policy makers should compare the method of APPROXI with the methods used in current models (LP, econometric, simulation). Since current models also require validation and since the APPROXI method tries to combine the strong aspects of these models, it is my opinion that the concept of APPROXI as described in this chapter requires further attention, investigation and development. An advantage of APPROXI is its flexibility and the ease to incorporate knowledge.

GENERAL DISCUSSION

"The farmer must take the ultimate decisions, and it's our task to provide the appropriate tools to support these decisions"

(Hennen and de Hoop, 1991)

Introduction

This thesis investigates the possibilities of methods and knowledge-based systems (KBSs) for the analysis of technical and economic accounting data from individual dairy farms to support their management and for use in models regarding policy evaluation.

Dairy farm management has become very important, and improvement is critical for a number of farms to survive. This requires a better analysis of available (accounting) data, and, since most farmers need to be supported for this task, KBSs might have potentials (chapter 1). Computerised analysis by a KBS requires not only accounting data, which are at the moment not uniform between different organisations, but also good standards for comparative analysis and especially knowledge. The standards currently used in the Netherlands are less suitable to be used in view of the criteria put forward for them (section 2.1). Therefore a new type of standard has been described and applied: the farm-adjusted standard or FAS (chapter 2). The concepts of KBSs are briefly and generally described, followed by a motivation of the approach in this thesis as different from the approaches found in literature (chapter 3). Two methods in the field of Artificial Intelligence have been developed and applied. The first one is IMAGINE, which can be used in quantitative domains characterised by a combinatorial explosion of possible situations (chapter 4). Smooth or fuzzy boundaries and the possibility of compensation between concepts is not only the essence of IMAGINE, but also of the second method which gives clearer explanation and which is also suitable for domains where data are qualitative and uncertain: FUZZY-DETECTOR (chapter 5). The methods IMAGINE and FUZZY-DETECTOR have been programmed as software products or tools (with the same names) for the development of KBSs. These tools have been used to develop GLOBAL-DETECTOR, a KBS for the analysis of gross margin on dairy farms (chapter 6), and ENVIRONMENT-DETECTOR, a KBS for analysis and planning to reduce nitrogen surplus while maintaining income (chapter 7). FASs are used in both KBSs. Finally, the APPROXI method is described and applied after a first model has been built based on this method. APPROXI estimates sector responses on a policy measure based on an individually used KBS (i.e. ENVIRONMENT-DETECTOR as example in chapter 8), as alternative of econometric and linear programming models.

In this final chapter, a general discussion is directed towards the degree to which the objective and requirements of this thesis are met, the prospects and limitations of the proposed methods and KBSs, users' involvement in developing systems and methods, stimulation of farmer's creativity, the attitude to computerised advices, and suggestions for a futurous research agenda. This chapter will close with a list of the main conclusions.

Objective of this thesis and accomplishment of requirements

The objective of this thesis has been the investigation of the possibilities for developing methods and KBSs for the analysis of technical and economic accounting data from individual dairy farms to support mainly the evaluation and tactical management functions, and for sector responses on policies. The study has shown that it is possible to develop methods and two KBSs for these tasks. For the investigation they were actually developed and tested, and described in this thesis. To assess the degree to which we have reached our objective, the developed methods and KBSs are matched with their requirements 1) as described in section 1.4.

1. Data from year-end accounts, which are already available and stored in the data base of the FADN, are used. Presence of the farmer or additional information is not required. The two KBSs cannot directly be linked with data bases from other accountancies, due to the disuniformity of the data. However, for two accountancies we have proved that simple data-transformation programmes in combination with little adaptations of the KBSs make this possible.
2. Both KBSs can give suggestions for improvement of the situation, and from our experiences from e.g. the farmers of the test group, we have reason to believe that they support the management within the current farm set-up.
3. Much emphasis is placed on accounting for the specific situation of the farm, especially by the use of FASs (De Haan 1991, and chapter 2). Since the actual individual farm specific input-output relations are not known under practical circumstances, they can of course not be used. But we have tried to make such relations quite farm specific by the use of FAS for a group of farmers (section 2.4.2). Although the KBSs were developed according to wishes and requirements of farmers who were involved during the whole development phase, our KBSs do neither intrinsically account for farmer's individ-

1) Many of these requirements are put forward as a result of the research by De Hoop et al. (1988) carried out on about two dozens of dairy farms.

ual need, wishes and style of farming, nor for different types of decision and information behaviour (Bemelmans, 1987). The method FUZZY-DETECTOR in this thesis has potentials to account for such factors. However, we did not account for them because (1) the KBSs, together with their extensive explanation facilities, are flexible enough for each farmer to use them according to his own wishes, etc, and (2) data are not available or there is effort needed to acquire them, and (3) farmers want to make their own corrections for individual wishes, style of management, etc; they want an objective (economic) comparison (section 6.5.5).

4. The use of FASs makes it possible to compare farm results with the results from other comparable farms. The differences between actual results and FASs values yield the building blocks for the knowledge-based generation of strong and weak aspects regarding the management and suggestions for improvement. Such results would not have been obtained when other, currently applied, standards were used. It is to be expected that FASs will emerge at organisations since they form a better basis for comparison than other standards, especially in the dairy sector (chapter 2).
5. Emphasis is on is the ability of our KBSs to give insight, and it is likely that farmers who use them will learn from them. Extensive but optional explanation facilities give the required (general) information. Farmers of the test group gave suggestions concerning explanation facilities and for more clearness. An example is that their remarks led amongst others to the development of FUZZY-DETECTOR (chapter 5), which is easier to understand than IMAGINE.
6. Both KBSs stimulate the creativity of farmers. The suggestions for improvement, which are given under objective and economic considerations, set farmers thinking although they may disagree with them. The creativity is also stimulated by the many options of GLOBAL-DETECTOR and the search for interesting alternatives in ENVIRONMENT-DETECTOR.
7. Maintenance of the KBSs is easy to perform, especially for GLOBAL-DETECTOR. Yearly updates of the coefficients for the FASs are automatically read in. IMAGINE and FUZZY-DETECTOR advocate fast development of the knowledge bases as well as a fast and easy maintenance of them.
8. Both KBSs advocate also widespread use. The KBSs can be used on the farm with no or only moderate support because of the user-friendliness (especially regarding GLOBAL-DETECTOR) and the moderate hardware requirements (a PC without extended memory and without auxiliary programmes). Since most farmers do not own a PC, the systems are also intended to be used at extension offices or on the portable PC of extension workers visiting farmers. GLOBAL-DETECTOR has a so-called 'coach version'. Widespread use is also advocated because the KBSs require only data that are already available at data bases. Manual data entry is neither necessary nor rec-

ommended, though possible. Finally, the KBSs can be used on an accountancy with direct access to the data base and (part of) the results of the analysis can be mailed to the farmers together with the report.

9. The KBSs presented in this chapter support dairy farm management. Data, knowledge, explanation texts and algorithms for the dairy domain are separated from the general control structures, user-interfaces and input-output facilities of these KBSs. Because of this separation it is possible to replace data, knowledge, etc, from the dairy domain with the data, knowledge, etc, from other domains. Early prototypes have been developed for arable, pig breeding and poultry farms by using the stripped GLOBAL-DETECTOR (i.e. use as a tool). In such a way similar systems can be developed for other domains, even outside agriculture. The same is true, of course, for the tools of IMAGINE and FUZZY-DETECTOR which are part of our KBSs. IMAGINE has been used several times, e.g. for the identification of styles of farming for APPROXI. We have seen in chapter 8 that the KBS ENVIRONMENT-DETECTOR is the central part of the APPROXI model to estimated sector responses on policy options regarding the nitrogen surplus.

From the way the results of the study match the requirements mentioned above, it can be concluded that the developed methods and KBSs have many opportunities to support management, to have better use of accounting data, and to support policy makers.

The two KBSs and the APPROXI model described in this thesis are merely aids for decision-making by farmers or policy makers. The systems are therefore not very detailed, but many aspects are brought together integrally. They intend to stimulate the creativity (see below).

Prospects and limitations of proposed methods and systems

For this study methods were developed in the field of Artificial Intelligence (IMAGINE and FUZZY-DETECTOR), which have been applied in two KBSs (GLOBAL- and ENVIRONMENT-DETECTOR). The motivation to use KBSs has been based on experiences reported in the literature (chapter 1). In this thesis it has not been investigated if 'conventional' systems (e.g. spread sheets, systems for optimalisation or simulation) would have been more suitable than KBSs. Although many authors are euphoric about KBSs, very few applications have proved to be successful. One cannot deny that 'conventional' systems also contain knowledge, and one might call many of them also KBSs. But the KBSs, or 'expert systems', in the sense they are meant and used in this study have some characteristics that make them different from 'conventional' systems, and these differences are also the very reason why KBSs have been chosen for analysis. Besides, the objective of this study was the investigation of the possibilities of such systems.

The characteristics that make KBSs so different from 'conventional systems', and that make them so suitable to be used for the analysis - according to the requirements (see above and section 1.4) - are the separation of knowledge from the control structures, the possibility to develop extended explanation facilities, and the use of heuristics from human experts for judgement (diagnosis). Compared to 'conventional' systems, the developed KBSs in this study are very flexible and changes and maintenance can be performed very easily. The flexibility regarding the handling of knowledge has been proved during the development of APPROXI when ENVIRONMENT-DETECTOR was extended. IMAGINE and FUZZY-DETECTOR have made acquisition, representation and maintenance of knowledge fast and easy, which supports our opinion that KBSs have great potentials for the analysis and interpretation of accounting data.

A limitation of the methods and KBSs is that they are developed with computer language LISP, which is rather uncommon in agricultural research. LISP has been chosen as language firstly to have a vehicle to get more acquainted in the field of Artificial Intelligence and secondly because some experience with LISP was required for the development of our first KBS with a so-called "empty shell" (Hennen, 1989).

There are some other limitations. Our methods and KBSs are not yet sufficiently validated and tested (more research is required), the KBSs are restricted to accounting data, the analysis is sometimes not detailed enough, and accounting offices and other organisations are unacquainted and very reserved regarding such new techniques.

Prospects are that accounting data will be used better and the management of the farmer can be supported. Without burdening with extra questions, the farmer is able to independently analyse his farm. Some farmers who tested both KBSs are quite enthousiastic. Remarks like "this is just what we needed" and "with such a system [GLOBAL-DETECTOR] you can earn more money than trundle a wheelbarrow" support our conviction that the KBSs have promises.

Earlier in this chapter it was remarked that the methods can be used in other domains. Hence, results from this research are not restricted to the dairy farm sector. Especially the Artificial Intelligence methods IMAGINE and FUZZY-DETECTOR may be valuable for developers of other KBSs.

User's involvement in developing systems and methods

As early prototypes, GLOBAL-DETECTOR and later ENVIRONMENT-DETECTOR have been installed on the PCs of six farmers from the test group. They used and tested it for a couple of hours and were asked to fill in a questionnaire about several aspects of the system. The answers were discussed in a meeting with these users, resulting in conclusions that directed further development. Such interactive development appeared to be very useful and important for the realisation of these KBSs.

The development of systems must fit the management behaviour and user's need to a satisfactory extent. The interactive or participative development of current systems for management support is generally too limited. More involvement of farmers and making use of the expertise of farmers will benefit not only the development of management information systems or KBSs, but also the development of methods (e.g. FUZZY-DETECTOR), and the development of models for policy support like APPROXI.

The tools of IMAGINE and FUZZY-DETECTOR are suitable 1) to support participative and interactive development since the acquisition and implementation of knowledge in a system with these tools is relatively easy.

Stimulation of farmer's creativity

The two KBSs may stimulate farmer's creativity because of their flexibility in use and because they explain how suggestions for improvement are inferred. This stimulation should be of permanent concern, and KBSs must not take over the creative process of management but they should 'merely' be an aid in this process. Besides, management is far more than the aspects covered by the KBSs.

The creativity may be stimulated further by asking the farmer to think thoroughly about his own mission, strategy, and tactics, and bring this into a system 2). The tools based on the methods described in this thesis may be flexible enough to be used. The farmer is supported in thinking about his mission, etc, and he can discuss this with an extension worker or with other farmers in a study group. Developers from management information systems might also benefit from such procedure since they can obtain an important source of information.

Computerised suggestions for tactical and strategic management: time will tell

GLOBAL- and ENVIRONMENT-DETECTOR yield suggestions for improvement of the management. The term 'advices' has not been used. Since we do not have all the available data and information of the farmer and his farm, and since we cannot overlook all the consequences when measures are applied, we would better use the term suggestions instead of advices. However, KBSs that are applied in operational

---

1) Especially after these tools are further developed to make automatic knowledge acquisition and maintenance possible (see below).

2) Boehlje and Eidman (1984) reported such approach that was called the production or service enterprise control system. Attonaty and Soler (1991) made a computer programme to construct a model of farmer's decision-making processes.

domains (e.g. pest management in horticulture) should stick to the term 'advices' because they use detailed information, the type of farmer and farming is far less important, and those KBSs are able to generate undisputable conclusions.

What we have noticed during several tests of GLOBAL-DETECTOR, especially at accountancies, is the reluctant attitude regarding the automatic generation of suggestions and the presentation of those to the farmer. Accountancies are presumably not ready for this, because most of them have not yet got a computerised analysis of individual farm records. Only when accountancies are accustomed to perform analysis for their clients, the next step might be computerised interpretation and generation of suggestions for improvement of tactical and strategic management. But for the moment it is uncertain how farmers will be advised in the future. The accountancies have the data and the extension workers and other technical advisors have the knowledge, so both parties gain from a better cooperation.

Future research

The following issues are proposed for future research:

1. Extensive validation and testing of methods and KBSs, and especially the APPROXI model. Not only by asking a judgement from other experts, but especially by means of questionnaires for farmers after they have tested the systems or after they are confronted with the results.
2. Further improvement of the APPROXI model, especially by including strategic aspects. The philosophy of APPROXI may also be used in the development of models for other kinds of policies or in other domains.
3. Extension of GLOBAL-DETECTOR with more possibilities for planning, so that the farmer obtains insight in the economic consequences of various measures he might take in future. Extension of ENVIRONMENT-DETECTOR with other environmental aspects might be of interest, because the current system only deals with the nitrogen surplus. There are possibilities to integrate both KBSs, e.g. to incorporate GLOBAL-DETECTOR's extensive facilities for analysis in ENVIRONMENT-DETECTOR.
4. Further development of IMAGINE (e.g. the use of other functions) and investigating the possibilities of this method for other applications to come to a method for general usage. This might be followed by the further development of IMAGINE to a tool for automatic knowledge acquisition and maintenance. With such a tool knowledge can be put in the knowledge base by the expert himself (independent from the knowledge engineer) in an easy way. The expert can eventually develop and maintain the knowledge base on his own.

5. Further development of FUZZY-DETECTOR, especially with more emphasis on the membership functions. Investigating the possibilities of this method and tool for other applications in order to come to a method and tool for general usage, eventually followed by the further development of FUZZY-DETECTOR to a tool for automatic knowledge acquisition and maintenance (like the one proposed for IMAGINE).
6. The development of comparable KBSs for other domains, based on or making use of the methods and tools described in this thesis. From GLOBAL-DETECTOR and from ENVIRONMENT-DETECTOR all domain dependent parts (i.e. knowledge, algorithms, etc, from the dairy domain) can be stripped off. What remains are two tools or 'empty shells' for the development of similar systems in other domains just by incorporating the domain knowledge, algorithms, etc, in these tools.
7. Exploring the value of FASs in other branches of agriculture.
8. Development of a system in which the farmer can bring his own mission, strategy and tactics as described above (stimulate creativity).

Main conclusions

1. The developed methods and KBSs for this thesis have opportunities to support management on dairy farms and to have a better use of accounting data.
2. For the analysis of farm results we must account for the specific situation of the farm. FASs are very suitable for this in the dairy sector, and without FASs the generation of strong and weak aspects regarding the management and suggestions for improvement would have been less easy and straightforward.
3. KBSs must make an objective and economic analysis. They must be flexible and transparant (with explanation facilities) so that farmers with various styles of farming and with different decision and information behaviour can use the same system.
4. The proposed method for sector responses on government policy measures (i.e. APPROXI), which is based on an individually used KBS (i.e. ENVIRONMENT-DETECTOR), is a good alternative for some econometric and linear programming models.
5. Both KBSs may have different kinds of users, and thereby a widespread use in the dairy farm sector with low required support and limited maintenance. They can also be used as a tool for the development of systems in other branches.
6. The development of KBSs need not be time-consuming when suitable methods and tools for the acquisition and representation of knowledge are used.
7. The methods and KBSs have many opportunities for application in other domains.

REFERENCES

Aarts, H.F.M., E.E. Biewinga and H.H. van Keulen (1992)
Dairy farming systems based on efficient nutrient management; In: Netherlands Journal of Agricultural Science, 40: 285-299

Abler, D.G., J.S. Shortle (1992)
Environmental and farm commodity policy linkages in the US and the EC; In: European Review of Agricultural Economists, 19: 197-217

AF, DLG, RASE (1992)
Farm computer technology in search of users; Proceedings of the third International Congress for Computer Technology in Agriculture; Paris - Versailles

Alleblas, J.T.W. (1988)
Management in de glastuinbouw, een zaak van passen en meten; Den Haag, Landbouw-Economisch Instituut; thesis (English summary)

Alleblas, J.T.W. (1991)
Towards an automated system of management advice in glasshouse horticulture; In: Acta Horticulturae nr. 295; Wageningen, ISHS

Attonaty, J.M. and L.G. Soler (1991)
Renewing strategic decision-making aids; In: Euro. R. agr. Eco., 18: 423-442

Baas, S.M. and H. Kwakernaak (1977)
Rating and ranking of multiple-aspect alternatives using fuzzy sets; In: Automatica, 13: 47-58

Baltussen, W.H.M. (1992)
Effectiviteit van stikstofheffingen voor landbouwbedrijven; Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in Dutch)

Baltussen, W.H.M., R. Hoste, C.H.G. Daatselaar and S.R.M. Janssens (1992)
Verschillen in mineralenoverschotten tussen bedrijven in de melkveehouderij en de akkerbouw; Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in Dutch)

Baltussen, W.H.M., D.W. de Hoop and J. van Os (1993a)
Economic instruments in agricultural policies to attain environmental goals; In: Proceedings of the VII European Congress of Agricultural Economics, entitled: Transition to an integrated agricultural economy. Stresa (Italy), september 6-10, 1993

Baltussen, W.H.M., J.A. Breembroek, E.R. ter Elst-Wahle and E.P.H. van de Ven (1993b)
Kasva knelpunten analyse systeem varkenshouderij. Onderzoek naar de mogelijkheden sterke en zwakke punten op te sporen van een bedrijf met varkens op basis van de financiële administratie; Rosmalen, Proefstation voor de Varkenshouderij (English summary)

Barrett, J.R., J.B. Morrison and L.F. Huggins (1985)
Artificial intelligence and expert systems in agricultural research and education; Paper, American Society of Agricultural Engineers; No. 85-5516

Barrett, J.R., T.L. Thompson and D.D. Jones (1990)
Knowledge systems development in U.S. agriculture; presentation at AgEng '90, International Conference on Agricultural Engineering, October 24-26, 1990, Berlin, Federal Republic of Germany

Bauer, S. (1989)
Historical review, experiences and perspectives in sector modelling; In: S. Bauer, W. Henrichsmeyer [eds.], Agricultural Sector Modelling, Proceedings 16th EAAE-Symposium; Vaug, Kiel

Becker, H. and H. Guyomard (1992)
Wirkungen von Preissenkungen und Düngemittelabgaben auf landwirtschaftliche Faktoreinkommen und Nährstoffbilanzen in der Europäischen Gemeinschaft; In: Agrarwirtschaft 41, Heft 10, 290-298

Bellman, R.E. and L.A. Zadeh (1970)
Decision-Making in a Fuzzy Environment; In: Management Science, vol. 17, pp. B-141-B-164

Bemelmans, T.M.A. (1987)
Bestuurlijke informatiesystemen en automatisering; Stenfert Kroese, Leiden/Antwerpen (in Dutch)

Berentsen, P.B.M. and G.W.J. Giesen (1993)
A comparison at farm level of different governmental policies to reduce N-losses in dairy farming; In: Proceedings of the VII European Congress of Agricultural Economics, entitled: Transition to an integrated agricultural economy; Stresa, Italy, september 6-10, 1993

Berentsen, P.B.M., G.W.J. Giesen and S.C. Verduyn (1992)
Manure legislation effects on income and on N, P, and K losses in dairy farming; In: Livestock Production Science, 31, 43-56

Berry, J.S., W.P. Kemp and J.A. Onsager (1991)
Integration of simulation models and an expert system for management of rangeland grasshoppers; In: AI Applications, 5, No.1, 1-14

Boehlje M.D. and V.R. Eidman (1984)
Farm Management; John Wiley & Sons, New York, etc

Boggess, W.G., P.J. van Blokland and S.D. Moss (1989)
FinARS: A financial analysis review expert system; In: Agricultural Systems, 31, 19-34

Bouwman, M.J. (1982)
The use of accounting information: expert versus novice behaviour; In: Ungson, G.R., D.N. Braunstein [eds.], Decision Making: an interdisciplinary inquiry; Kent publishing Company, Boston, Mass, USA

Bree, D.S. and W.H.G.J. Hennen (1989)
The use of expert systems for the interpretation of technical and economical performance of dairy farms; In: G. Schiefer [ed.], Expertensysteme in der Agrarwirtschaft: Entwicklung, Erfahrung, Perspectiven; Wissenschaftsverlag Vauk, Kiel

Breembroek, J.A. (1991)
KASA - Knelpunten Analyse Systeem Akkerbouw;
In: W.H.G.J. Hennen, et al. [eds.], Informatica toepassingen in de agrarische sector. Voordrachten VIAS-symposium 1991;
Agro-informaticareeks nr. 5 (in Dutch)

Breuker, J.A. and B.J. Wielinga (1984)
Elicitation techniques; In: Breuker, J.A., B.J. Wielinga [eds.], Techniques for knowledge elicitation and analysis; Report 1.5, Esprit project 12

Breuker, J.A. and B.J. Wielinga (1989)
Models of expertise in knowledge acquisition; In: Guida, G., C. Tasso [eds.], Topics in expert system design: methodologies and tools; Amsterdam, North Holland Publishing Company

Buchanan, B.G. and E.H. Shortliffe (1984)
Rule-based expert systems; Reading, MA: Addison-Wesley

Bullen, C.V. and J.R. Rockart (1981)
A primer on critical success factors; In: CISR No. 69, Center for Information Systems Research, Sloan school of Management; Massachusetts Institute of Technology

Burrell, A. (1989)
The demand for fertiliser in the United Kingdom; In: Journal of Agricultural Economics, 40(1), 1-20

Carrascal, M.J. and L.F. Pau (1992)
A survey of expert systems in agriculture and food processing; In: AI Applications, 6, No.2

CLM, DLV, IKC (1991)
Mineralenboekhouding Melkveehouderij; Utrecht (in Dutch)

Coughlan, J.C. and S.W. Running (1989)
An expert system to aggregate biophysical attributes of a forested landscape within a geographic information system; In: AI Applications, 3, No.4, 35-43

Daatselaar, C.H.G. (1988)
De invloed van de melkgift op het saldo per hectare; Den Haag, Landbouw-Economisch Instituut (in Dutch)

Daatselaar, C.H.G. (1989)
Verschillen in de mineralenbalans tussen melkveebedrijven. Een vergelijkend onderzoek; Den Haag, Landbouw-Economisch Instituut (in Dutch)

Daatselaar, C.H.G., D.W. de Hoop, H. Prins and B.W. Zaalmink (1990)
Bedrijfsvergelijkend onderzoek naar de benutting van mineralen op melkveebedrijven; Den Haag, Landbouw-Economisch Instituut (in Dutch)

Daatselaar, C.H.G., W.H.G.J. Hennen, D.W. de Hoop and B.W. Zaalmink (1993)
Optimaal produceren; Praktijkreeks Veehouderij; Misset, Doetinchem (in Dutch)

Davis, G.B. and M.H. Olson (1985)
Management Information Systems: Conceptual Foundations, Structure, and Development; New York, McGraw-Hill Book Company

Dijk, J.P.M. van, B.E. Douma and A.L.J. van Vliet (1994)
Bedrijfsuitkomsten in de landbouw (BUL); Periodieke Rapportage 11-92/93, Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in Dutch)

Dobbins, C.L. and R.P. King (1988)
An expert system for interpreting the year-end business summary of crop-hog farms; University of Minnesota, USA, Draft 2, July 1

Dobbins, C. (1989)
A review of farm business analysis research; Agricultural Economics Staff Paper # 89-8; Purdue University, West Lafayette, Indiana

Doluschitz, R. and W.E. Schmisseur (1988)
Expert systems: applications to agriculture and farm management; In: Computers and Electronics in Agriculture, 2, 173-182

Doluschitz, R. (1989)
Expertensysteme zum Einsatz im Management der Milcherzeugung; In: G. Schiefer [ed.], Expertensysteme in de Agrarwirtschaft: Entwicklung, Erfahrung, Perspectiven, Vortraege der ITMA-Fachtagung 'Expertensysteme 1988'. BRD, VAUK Kiel

Dombi, J. (1990)
Membership function as an evaluation; In: Fuzzy Sets and Systems, 35, 1-21

Elhorst, J.P. and H.G. van der Meer (1993)
Verkennende studie over input-output relaties: het effect van beperking van het verbruik van inputs op de produktie-omvang en emissies; Den Haag, NRLO (in Dutch)

Ericsson, K.A. and H.A. Simon (1984)
Protocol analysis: verbal reports as data; MIT Press, Cambridge, Massachussets

Evans, M., R. Mondor and D. Flaten (1989)
Expert systems and farm management; In: Canadian Journal of Agricultural Economics, 37, 639-666

Feigenbaum, E.A. (1977)
The art of artificial intelligence; themes and case studies of knowledge engineering; In: Proceedings IJCAI 77, Cambridge MA, Morgan Kaufman

Fillatre, C. and P.A. Moreau (1991)
COFINE: The initial large-scale of an expert system in the field of agricultural management; In: Proceedings of the XXVI EAAE Seminar 'Economics and Artificial Intelligence in Agriculture' 10-12 September, Grignon, France

Folkerts, H. (1989)
Expertsystemen in de rundveehouderij; In: Landbouwmechanisatie 2, 60-61 (in Dutch)

Fontein, P.F., G.J. Thijssen, J.R. Magnus and J. Dijk (1992)
Effect van regulerende heffingen ter reductie van het mineralenoverschot in Nederland; Tilburg, Economisch Instituut (in Dutch)

Freksa, C. (1982)
Linguistic description of human judgements in expert systems and the soft sciences; In: M. Gupta, E. Sanchez [eds.], Approximate reasoning in decision analysis; Amsterdam, North-Holland

French, S. (1984)
Fuzzy decision analysis: some criticisms; In: H.-J. Zimmermann, L.A. Zadeh, B.R. Gaines [eds.], Fuzzy sets and decision analysis; Elsevier Science Publishers B.V. (North-Holland)

Gaines, B.R., L.A. Zadeh and H.J. Zimmermann (1984)
Fuzzy sets and decision analysis - a perspective;
In: H.J. Zimmermann, L.A. Zadeh, B.R. Gaines [eds.], Fuzzy sets and decision analysis; Elsevier Science Publishers B.V. (North-Holland)

Gekle, L. (1990)
Erfolgsunterschiede landwirtschaftlicher Betriebe - untersucht anhand von Erfolgsgruppen der Buchfuerungsstatistik; In: Berichte aus der Landwirtschaft, 68, 82-100

Gold, H.J., G.G. Wilkerson, Y. Yu and R.E. Stinner (1990)
Decision analysis as a tool for integrating simulation with expert systems when risk and uncertainty are important; In: Comput. Electron. Agric., 4, 343-360

Gordon, S.E., R.T. Gill and T.A. Dingus (1987)
Designing for the user: the role of human factors in expert system development; In: AI Applications, 1, No.1, 35-45

Gordon, S.E. (1989)
Theory and methods for knowledge acquisition; In: AI Applications, 3, No.3, 19-30

Graham, I. (1991)
Fuzzy logic in commercial expert systems - Results and prospects; In: Fuzzy Sets and Systems, 40, 451-472

Haan, T. de (1990)
Unpublished comments concerning research on hetero-scedasticiteit in farm-adjusted standards

Haan, T. de (1991)
Het ontwikkelen van bedrijfsspecifieke vergelijkingsmaatstaven voor de analyse van het bedrijfsresultaat op melkveebedrijven; Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in Dutch)

Haan, T. de (1995)
Rekenregels voor de effecten van diverse voedertaktieken op de melkproduktie en het bedrijfsresultaat; Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in Dutch) (in press)

Harrison, S.R. (1991)
Validation of Agricultural Expert Systems; In: Agricultural Systems, 35, 265-285

Harsh, S.B. (1988)
The role of expert systems as an extension tool for management education; Agricultural Economics Staff Paper # 88-42, Michigan State University, East Lansing

Harsh, S.B. (1990)
Future role of computers in your farm business; Agricultural Economics Staff Paper # 90-74, Michigan State University, East Lansing

Harsh, S.B. (1991)
Modelbase structure and design; Handout for post graduate course 'Integrated Decision Support Systems (IDSS) in farm management' at Department of Agricultural Economics, University of Wageningen

Hayes-Roth, F., D.A. Waterman and D.B. Lenat (1983)
Building Expert Systems; Addison-Wesley Publishing Company, Wokingham, England, etc

Heinrich, I. and K. Walter (1989)
Kontrolle und Analyse des Betriebsweiges Futterbau - Rindviehaltung mit Personalcomputer; In: Agrarwirtschaft, 38, Heft 5, 135-150

Hennen, W.H.G.J. (1989)
De methode voor het bouwen van expert systemen. Getest voor de analyse van de omzet op melkveebedrijven; Den Haag, Landbouw-Economisch Instituut (in Dutch)

Hennen, W.H.G.J. (1991)
Methoden en technieken voor kennis-elicitatie ten behoeve van de ontwikkeling van expert systemen in het B2-project van SKBS; Uitgave SKBS, Erasmus Universiteit Rotterdam (in Dutch)

Hennen, W.H.G.J. and D.W. de Hoop (1991)
GLOBAL-DETECTOR: knowledge-based analysis and diagnosis of economical performance on dairy farms; In: Proceedings of the 26th EAAE Seminar 'Economics and artificial intelligence in agriculture' 10-12 September, Grignon, France

Hennen, W.H.G.J. and J.J.F. Wien (1994)
Documentatie Milieu-Detector; Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in Dutch) (in press)

Hoop, D.W. de (1988)
Management processes in dairy and pig farming and the construction of systems; In: Modelling of livestock production systems, Seminar in the European Community Programme for the Coordination of Agricultural Research, held in Brussels, Belgium, 7-9 April 1987, pp. 77-86

Hoop, D.W. de (1991)
Ontwikkeling en gebruik van informatiesystemen voor de tactische planning en bedrijfsanalyse; In: Bedrijfsanalyse en -beheer, PHLO-cursus, Wageningen (in Dutch)

Hoop, D.W. de, J. Engelsma and G.J. Wisselink (1988)
De tactische boer: het management en de informatiebehoefte van melkveehouders; Den Haag, Landbouw-Economisch Instituut (in Dutch)

Huirne, R.B.M. (1990)
Computerized management support for swine breeding farms; Thesis, Department of Farm Management, Wageningen Agriculture University, Wageningen

Jackson, P. (1986)
Introduction to expert systems; Addison-Wesley Publishing Company, Wokingham, England, etc

Jain, R. (1977)
A procedure for multiple-aspect decision making using fuzzy sets; In: Int. J. Systems Sci., 8, no. 1, 1-7

James, S.C. and E. Stoneberg (1986)
Farm Accounting and Business Analysis; Ames, Iowa State University Press

Janssen, R. (1991)
Multiobjective decision support for environmental problems; Thesis, Amsterdam University (VU), Elinkwijk

Johnson, P.E. (1983)
What kind of expert should a system be?; In: J. Med. Philos., 8, 77-97

Kickert, W.J.M. (1978)
Fuzzy theories on decision-making; Martinus Nijhoff Social Sciences Division, Leiden/Boston/London

King, R.P. and S.B. Harsh (1986)
A preliminary proposal for a regional research project on farm information systems; Revision of a proposal presented at NCR-113 meeting, October 8, 1986, St Louis, Missouri

King, R.P., S.B. Harsh and C.L. Dobbins (1990)
Farm information systems: farmers needs and system design strategies; In: Tijdschrift voor Sociaal wetenschappelijk onderzoek van de Landbouw (TSL), 5 (1), 34-59

Klink, J.P.M. (1991)
Hoofdlijnen post-INSP-beleid: van informaticastimulering naar informatiebeleid; Gouda, Agritect Advies (in Dutch)

Kwakernaak, H. (1979)
An algorithm for rating multiple-aspect alternatives using fuzzy sets; In: Automatica, 15, 615-616

Lane, P.W., N.W. Galwey and N.G. Alvey (1988)
Genstat 5: an introduction; Oxford University Press

Leeuwis, C. (1993)
Of computers, myths and modelling: the social construction of diversity, knowledge, information, and communication technologies in Dutch horticulture and agricultural extension; Thesis, Wageningen Agricultural University

Leneman, H. (1993)
Schatting relatie bedrijfskenmerken op melkveebedrijven met stikstofemissie; Intern rapport, Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in Dutch)

LEI/VLB (1989)
Gedetaileerd informatiemodel 'analyse en diagnose'; Den Haag/Leiden, LEI/VLB (in Dutch)

Lindsay, P.H. and D.A. Norman (1977)
Human Information Processing; Academic Press, London

Longchamp, J.Y., J.P. Nicoletti and L. Lescar (1990)
A Coherent set of Decision Support Software for Farm Management; In: Proceedings of the 3rd international DLG-Congress for computer technology on Integrated Decision Support Systems in Agriculture. BRD, Frankfurt a.M.

Luger, G.F. and W.A. Stubblefield (1989)
Artificial Intelligence and the design of expert systems; Benjamin/Cummings, Redwood City, California

McGrann, J.M., K. Karkosh and C. Osborne (1989)
Agricultural financial analysis expert systems: software description; In: Canadian Journal of Agricultural Economics, 37, 695-708

Morahan, A.G., D.J. Ruane and M. Butler (1989)
The development and evaluation of a microcomputer based management information system for a dairying enterprise; In: V.A. Dodd, P.M. Grace [eds.], Proceedings of the 11th international congress on agricultural engineering [CIGR], Dublin, Ireland, 4-8 September 1989, 2761-2767

Nagaki, M. (1992)
Computer aided dairy farm management decision-making in Japan; In: Proceedings of the 4th international Congress for Computer Technology in Agriculture on Farm Computer Technology in Search for Users, Paris-Versailles, France, June, 111-114

Negoita, C.V. (1985)
Expert systems and fuzzy systems; The Benjamin/Cummings Publishing Company Inc., California

NRLO (1991)
Informatietechnologie in de Land- en Tuinbouw; NRLO-rapport 91/3, Den Haag, Nationale Raad voor Landbouwkundig Onderzoek (in Dutch)

Nuthall, P.L. (1992)
Successful development, marketing and support of farm financial software; In: Proceedings of the 4th international Congress for Computer Technology in Agriculture on Farm Computer Technology in Search for Users, Paris-Versailles, France, June, 121-124

O'Keefe, R.M., O. Balci and E.P. Smith (1988)
Validating expert system performance; In: AI Applications, Vol.2, No.2-3, 35-43

Palmer, G. (1992)
The use and application of computer technology on the farm: a first hand few; In: Proceedings of the 4th international Congress for Computer Technology in Agriculture on Farm Computer Technology in Search for Users, Paris-Versailles, France, June, 9-14

Phillips, J.J. and S.B. Harsh (1987)
An expert system application to the financial analysis of lender case farm records; Agricultural Economics Staff Paper # 87-69, Michigan State University, East Lansing

Ploeg, J.D. van der (1993)
Bedrijfsstijlen en classificatieschema's: over de aard van agrarische ontwikkelingsprocessen; In: Tijdschrift voor Sociaal wetenschappelijk onderzoek van de Landbouw (TSL), 8 (1), 50-70 (in Dutch)

Ploeg, J.D. van der, S. Miedema, D. Roep, R. van Broekhuizen and R. de Bruin (1992)
Boer Bliuwe, Blinder...!; AVM/CCLB en Vakgroep Agrarische Ontwikkelingssociologie van de Landbouwuniversiteit, Wageningen (in Dutch)

Poppe, K.J. (1991)
Using Farm Accounts: a survey and recommendations for further research; In: K.J. Poppe [ed.], Information needs and accounting in agriculture, Den Haag, Landbouw-Economisch Instituut (LEI-DLO)

Poppe, K.J. (1992)
Accounting and the environment; In: G. Schiefer [ed.], Integrated Systems in Agricultural Informatics, Bonn, Germany

Quartel, E., G.J. Verschoor and P.M.J. Warffemius (1992)
Organisatie vs Expert Systeem. Sociale aspecten bij de invoering van Global Detector nader bekeken; Project Kennissystemen & Kennismanagement, Erasmus University Rotterdam, Fac. Management Science (in Dutch)

Roep, D., J.D. van der Ploeg and C. Leeuwis (1991)
Zicht op duurzaamheid en kontinuïteit, bedrijfsstijlen in de Achterhoek; Vakgr. Agr. Ontwikkelingssociologie, Wageningen (in Dutch)

Schakenraad, M.H.W., W.H.G.J. Hennen and D.W. de Hoop (1992)
The feed and grassland module: an integrated management support system for analysis of the feed and grassland management; In: Proceedings of the SKBS/MPP Symposium, Zeist

Schakenraad, M.H.W., W.H.G.J. Hennen and D.W. de Hoop (1994)
Reference management and conditional causal modelling to support feed and grassland management on Dutch dairy farms; Research-report. Den Haag, Landbouw-Economisch Instituut (LEI-DLO) (in press)

Smith, T.R. (1989)
The potential application of expert systems in dairy extension education; In: Journal of Dairy Science, 72 (10), 2760-2766

Steele, G.L. Jr. (1994)
Common LISP: The Language; USA, Digital Press

Stone, N.D. (1989)
Knowledge-based systems as unifying paradigm for IPM; In: Proceedings National IPM Symposium/Workshop, pp. 13-24, Las Vegas, Nevada

Swaans Arons, H. de, and P. van Lith (1984)
Expert systemen; Academic Servive, Den Haag (in Dutch)

Theil, H. (1979)
Principles of econometrics; Wiley & Sons, Inc., Amsterdam/New York/Oxford

Veen, M. van der, and J. Dijk (1993)
The nutrient flow model for Dutch agriculture: a new tool for environmental policy evaluation; In: Proceedings of the VII European Congress of Agricultural Economics, entitled: Transition to an integrated agricultural economy. Stresa (Italy), september 6-10, 1993

Wagner, P. (1992)
Methodische Grundlagen und praktische Entwicklung eines Expertensystems fuer die Wirtschaftlichkeitsanalyse landwirtschaftlicher Betriebe; Agrarwirtschaft, 132

Wagner, P. (1993)
Techniques of representing knowledge in knowledge-based systems; In: Agricultural Systems, 41, 53-76

Wain, N., C.D.F. Miller and R.H. Davis (1988)
Rule-based inference system for animal production management; In: Computers and Electronics in Agriculture, 2, 277-300

Waterman, D.A. (1986)
A guide to Expert Systems; Reading, etc, Addison-Wesley Publishing Company

Webster, J.P.G. (1988)
Applications in Management and Marketing; In: Proceedings of the 2nd International Congress for Computer Technology on Knowledge Based Systems in Agriculture, DLG, Frankfurt a.M.

Webster, J.P.G. and J.J. Amos (1987)
Expert systems for agricultural management; In: Farm Management, 24

Whalen, T. and B. Scott (1983)
Decision support with fuzzy production systems; In: P.Wang [ed.], Advances in fuzzy sets, possibility theory, and applications; Decision Science Laboratory, Georgia State University, Atlanta, New York:Plenum

Winston, P.H. (1984)
Artificial Intelligence; Addison-Wesley, Reading etc

Wossink, G.A.A. (1993)
Analysis of future agricultural change. A farm economics approach applied to Dutch arable farming; Thesis, Wageningen Agricultural University

Zaalmink, B.W., A. van der Kamp and J.J.F. Wien (1991)
Tact-melkvee: een beslissingsondersteunend model voor de melkveehouder; In: W.H.G.J. Hennen, et al. [eds.], Informatica Toepassingen in de Agrarische Sector, Voordrachten VIAS - Symposium 1991, Agro-informatica reeks, 4 (in Dutch)

Zachariasse, L.C. (1974)
Boer en bedrijfsresultaten. Analyse van uiteenlopende rentabiliteit van vergelijkbare bedrijven in de Noord-Oost polder; Thesis, Publikatie no. 8, Afdeling voor Agrarische Bedrijfseconomie aan de Landbouwhogeschool Wageningen (English summary)

Zachariasse, L.C. (1990)
Managementcapaciteiten: elementen en beïnvloeding; In: Achtergrond en gebruik van economische modellen en expertsystemen ter ondersteuning van managementbeslissingen op melkvee- en varkensbedrijven, PHLO-cursus, Wageningen (in Dutch)

Zachariasse, L.C. (1991)
Farm management information systems planning and development in the Netherlands; Paper presented at symposium 'harmonization of information systems: a farm business model approach', American Agricultural Economics Associations Meeting 1991, Manhattan, aug 1991

Zadeh, L.A. (1965)
Fuzzy sets; In: Information and control 8, 338-535

Zadeh, L.A. (1973)
Outline of a new Approach to the analysis of complex systems and decision process; In: IEEE Trans. Systems, Man, and Cybernetics, SMC-3(1973), 28-44

Zadeh, L.A. (1983)
The role of fuzzy logic in the management of uncertainty in expert systems; In: Fuzzy Sets and Systems, 11, 199-227

Zimmermann, H.J. (1987)
Fuzzy sets, decision making and expert systems; Kluwer Academic Publishers, Boston, Dortrecht, Lancaster

Zimmermann, H.J. (1991)
Fuzzy set theory and its applications; Kluwer Academic Publishers. Boston, Dortrecht, Lancaster

CURRICULUM VITAE

Willem Hubert Gemma Johannes Hennen werd geboren op 12 juni 1959 te Heerlen, en groeide op te Nieuwstadt. Hij behaalde achtereenvolgens de diploma's voor MAVO (Mr. Nic. Beckers, Sittard), HAVO en Atheneum-B (Bisschoppelijk College, Sittard). In 1979 begon hij met de studie Zoötechniek aan de Landbouwhogeschool te Wageningen. Het doctoraalexamen werd afgelegd in 1986 met als hoofdvak Veeteelt en als bijvakken Industriële Bedrijfskunde, Agrarische Bedrijfseconomie en Gezondheids- en Ziekteleer der Huisdieren.

In februari 1986 trad hij in dienst bij het Centrum voor Onderwijs in de Dierveredeling en de Rundveehouderij te Horst als docent economie. November 1986 stapte hij over naar het Landbouw-Economisch Instituut (LEI-DLO) waar hij - in het kader van het INSP - met name onderzoek deed naar de mogelijkheden van kennissystemen voor de analyse van boekhoudgegevens van rundveebedrijven. De belangrijkste methoden en kennissystemen uit de "DETECTOR-familie" die hij - met hulp van collega's uit de Sectie Veehouderij - heeft ontwikkeld, zijn in dit proefschrift beschreven.