

RIVM report 550002005/2004

**The MAP COMPARISON KIT:  
methods, software and applications**

H. Visser (editor)

This investigation has been performed by order and for the account of RIVM, within the framework of project S/550002/01/TO, Tools for Uncertainty Analysis.

## Abstract

Comparing maps is an important issue in environmental research. Reasons for comparing maps may be: (i) the different socio-economic scenarios on which they are based, (ii) detection of temporal changes, (iii) calibration/validation of land-use models, (iv) hot-spot detection, (v) their use in uncertainty analysis, and (vi) their origin in different methodologies/models. This report addresses the problem of quantifying subsequent map similarities and dissimilarities.

Our main focus is on maps denoted as ‘categorical’ or ‘nominal’. A number of the five map-comparison techniques are described. These techniques differ in mathematical approach (no math, ‘cell by cell’, two types of ‘fuzzy’ and ‘single-map statistics’) and apply to different types of maps (nominal, ordinal, ratio and interval scale). Special attention is given to the comparison of maps through *fuzzy-set calculation rules*. The rationale is that fuzzy-set map comparison is very close to human judgement, as shown in an Internet experiment.

The MAP COMPARISON KIT (MCK) software plays a major role in the report. MCK, a software package for ‘state-of-the-art’ map comparison, contains all the examples used in this report. The software, developed by order of the Netherlands Environmental Assessment Agency, was fully designed by the Research Institute for Knowledge Systems. The software will be made publicly available on the RIVM website early 2004 ([www.rivm.nl](http://www.rivm.nl)).

## Acknowledgements

A number of individuals and institutes have contributed to this report. First of all, let me thank the staff at RIKS, the Research Institute for Knowledge Systems. Inge Uljee and Alex Hagen did a marvellous job in designing and programming the MAP COMPARISON KIT software. In good cooperation with RIVM, they improved the various versions of the software needed to realize the final version. Alex Hagen also (co-)authored sections 2.3 and 2.5, and Appendix C on fuzzy-set map comparison versus human judgement. He was the initiator of section 4.1, the analysis of children's maps. Many thanks to Guy Engelen for both his encouragement and his coordination of the software development at RIKS.

A number of RIVM colleagues have made the following contributions: Ton de Nijs (RIM/MNP) as co-author of Chapter 1 and section 5.2, and lead author of section 5.1, on the calibration of the LUMOS Environment Explorer; Raymond de Niet (RIM/MNP) as lead author of section 5.2 on the consistency of land-use maps, and Kees Klein Goldewijk (KMD/MNP) as author of section 4.2 on the analysis of historical agricultural land use. Furthermore, Michel Bakkenes (NLB/MNP) provided the maps on heavy metals in topsoils (section 4.6); Laurens Zwakhals (VTV) the maps on ambulance driving times (section 4.5) and Wideke Boersma (RIM/MNP), a number of maps from the LUMOS Land Use Scanner (sections 4.3 and 4.4). Judith Borsboom-van Beurden (RIM/MNP) was co-author of sections 4.3 and 4.4.

Kit Buurman (IMP/MNP) played an important role in the initial phase of the Map Comparison Project. I thank Geert Verspaj and Gert-Jan Stolwijk (both IMP/MNP) for programming and testing a prototype user shell of the MAP COMPARISON MODULE within ArcGis-8. Finally, Arthur Beusen and Arnold Dekkers (both IMP/MNP) designed software tools additional to the MAP COMPARISON KIT software.

Peter Verburg (Wageningen UR) was the lead author of Appendix A on single-map statistics. The programming of enrichment factors was performed by Kor de Jong (University of Utrecht). He was also the co-author of Appendix A.

The software development was funded by several project leaders at the RIVM, i.e.: Arthur Beusen (IMP/MNP), Harm van den Heiligenberg (NMD/MNP), Peter Janssen (IMP/MNP), and Marianne Kuypers-Linde (RIM/MNP).

Hans Visser



# Contents

<b>SAMENVATTING .....</b>	<b>8</b>
<b>SUMMARY .....</b>	<b>9</b>
<b>1. INTRODUCTION.....</b>	<b>11</b>
1.1 THE ROLE OF MAPS IN ENVIRONMENTAL RESEARCH .....	11
1.2 WHY COMPARE MAPS? .....	12
1.3 FROM CELL-BY-CELL TO FUZZY.....	15
1.4 THE MAP COMPARISON KIT .....	17
1.5 THIS REPORT .....	17
<b>2. MAP COMPARISON - METHODS .....</b>	<b>19</b>
2.1 DEFINITION OF MEASUREMENT SCALES .....	20
2.2 VISUAL MAP COMPARISON .....	22
2.3 COMPARING MAPS CELL BY CELL - KAPPA STATISTICS.....	25
2.3.1 <i>Equal or unequal, that's the question</i> .....	25
2.3.2 <i>Advanced use of Kappa statistics</i> .....	25
2.3.3 <i>Kappa dissected into K-histo and K-location</i> .....	28
2.3.4 <i>Example</i> .....	30
2.4 FUZZY-SET MAP COMPARISON .....	32
2.4.1 <i>Considering categorical similarity</i> .....	32
2.4.2 <i>Considering proximity of similar cells</i> .....	33
2.4.3 <i>Comparison of fuzzy cells</i> .....	33
2.4.4 <i>Aggregated map results and overall similarity with Fuzzy Kappa</i> .....	35
2.4.5 <i>Dublin example continued</i> .....	35
2.5 COMPARING MAPS BY HIERARCHICAL FUZZY PATTERN MATCHING.....	38
2.6 COMPARING MAPS BY USE OF SINGLE MAP STATISTICS.....	39
<b>3. MAP COMPARISON - SOFTWARE .....</b>	<b>41</b>
3.1 MAP COMPARISON KIT.....	41
3.1.1 <i>The software</i> .....	41
3.1.2 <i>Handling different measurement scales</i> .....	43
3.2 TOOLS ADDITIONAL TO THE MAP COMPARISON KIT .....	46
3.3 SOFTWARE FOR THE CALCULATION OF ADDITIONAL STATISTICS .....	46

<b>4. MAP COMPARISON KIT – APPLICATIONS .....</b>	<b>49</b>
4.1 SPOT THE 15 DIFFERENCES! .....	50
4.1.1 <i>Introduction</i> .....	50
4.1.2 <i>Cell-by-cell comparison</i> .....	51
4.1.3 <i>Fuzzy-set map comparison</i> .....	53
4.1.4 <i>Conclusion</i> .....	55
4.2 HISTORICAL PROJECTIONS OF AGRICULTURAL LAND USE, 1700-1990 .....	56
4.2.1 <i>Context</i> .....	56
4.2.2 <i>Data</i> .....	56
4.2.3 <i>Similarity based on Kappa statistics</i> .....	58
4.2.4 <i>Fuzzy-set map comparison</i> .....	62
4.2.5 <i>Effect of aggregation</i> .....	64
4.2.6 <i>Conclusion</i> .....	66
4.3 THE LAND USE SCANNER AND SCENARIOS FOR THE YEAR 2020 .....	67
4.3.1 <i>The Land Use Scanner</i> .....	67
4.3.2 <i>About the model</i> .....	68
4.3.3 <i>Scenarios</i> .....	69
4.3.4 <i>Comparing the different scenarios</i> .....	70
4.4 RESIDENTIAL LAND USE IN 2020 .....	74
4.5 TRAVEL TIMES OF AMBULANCES .....	78
4.6 HEAVY METALS IN TOPSOILS .....	81
<b>5. MAP COMPARISON KIT - SPECIAL TOPICS.....</b>	<b>85</b>
5.1 CALIBRATION OF LAND-USE MODELS.....	85
5.1.1 <i>Method</i> .....	85
5.1.2 <i>Environment Explorer</i> .....	85
5.1.3 <i>Calibration of the Environment Explorer for 1996</i> .....	89
5.2 CONSISTENCY IN LAND-USE MAPS .....	93
5.2.1 <i>Context</i> .....	93
5.2.2 <i>From CBS land-use maps to LUMOS land-use maps</i> .....	94
5.2.3 <i>Locating and identifying suspect changes by the MCK</i> .....	94
5.2.4 <i>Fuzzy-set map comparison</i> .....	95
5.2.5 <i>Categories in detail</i> .....	97
5.2.6 <i>Formulating correction rules</i> .....	99
5.2.7 <i>Conclusions</i> .....	101
<b>6. CONCLUSIONS AND FURTHER RESEARCH.....</b>	<b>103</b>
<b>REFERENCES.....</b>	<b>107</b>
<b>APPENDIX A NEIGHBOURHOOD CHARACTERISTICS.....</b>	<b>113</b>
<b>APPENDIX B MAP TRANSFORMATIONS.....</b>	<b>120</b>
<b>APPENDIX C HUMAN JUDGEMENT AND FUZZY-SET MAP COMPARISON.....</b>	<b>123</b>



## Samenvatting

Kaarten, en in het bijzonder landgebruikskarten, worden gepresenteerd in een reeks van uitgaves van het Milieu- en Natuurplanbureau (MNP/RIVM). Voorbeelden zijn de Milieu- en Natuurverkenning die eens per vier jaar uitkomen, als ook de Milieubalans en Natuurbalans die jaarlijks worden gepubliceerd. Kaarten en kaartbeeldvergelijkingen spelen een belangrijke rol bij het tot stand komen van deze producten. Hierbij kunnen kaarten voor allerlei doeleinden vergeleken worden: (i) kaarten kunnen gebaseerd zijn op verschillende economische/demografische scenario's, (ii) detectie van veranderingen in landgebruik in de tijd, (iii) calibratie en validatie van landgebruiksmodellen, (iv) detectie van hot-spots in kaarten, (v) onzekerheidsanalyse en (vi) analyse van kaarten die afkomstig zijn van verschillende onderzoeksgroepen of modellen.

Dit rapport is gericht op het kwantificeren van verschillen en overeenkomsten in kaarten. Hierbij ligt de nadruk op nominale en ordinale kaarten. Voor dit type kaarten is het uitvoeren van vergelijkingen het lastigst. Immers, hoe zou je verschillen tussen de nominale landgebruikscategorieën 'grasland', 'bebouwing' of 'recreatie' numeriek moeten uitdrukken?

We beschrijven een vijftal methodes om kaartbeelden te vergelijken. Deze methodes verschillen in wiskundige benadering (geen wiskunde, 'cel bij cel', twee soorten 'fuzzy' en kengetallen voor enkelvoudige kaarten) als ook in het type kaarten waarvoor ze bedoeld zijn (nominale schaal, ordinale schaal, ratio-schaal en interval-schaal). Speciale aandacht wordt besteed aan kaartbeeldvergelijking gebaseerd op rekenregels uit de fuzzy-set theorie. Het sterke punt van deze benadering is dat door het fuzzificeren van categorieën en hun specifieke locatie, een methode ontstaat die veel overeenkomsten bezit met hoe het menselijk oog kaarten vergelijkt. We laten dit zien aan de hand van een Internet-experiment.

De kaartbeeldvergelijkingsmethodes worden geïllustreerd aan de hand van een aantal case-studies. De voorbeelden variëren van de analyse van historisch agrarisch grondgebruik over de periode 1700 tot 1990, tot het analyseren voor landgebruikstoekomstbeelden voor het jaar 2020. Speciale toepassingen zijn het calibreren en valideren van landgebruiksmodellen, en het zoeken naar inconsistenties in landgebruikskarten van het CBS.

Alle analyses uit het rapport zijn uitgevoerd met de MAPCOMPARISONKIT-software. Dit software pakket is in opdracht van het RIVM ontwikkeld door het Research Instituut voor KennisSystemen (RIKS). De software is uniek in het integreren van vier kaartbeeldvergelijkingsmethodes. De software zal begin 2004 algemeen beschikbaar worden gesteld via de RIVM-website ([www.rivm.nl](http://www.rivm.nl)). We concluderen dat de nieuw ontwikkelde MAPCOMPARISONKIT-software een krachtig 'state of the art' hulpmiddel is bij het analyseren van kaartbeelden.

## Summary

The use of (land-use) maps is important for a wide range of products issued by the Netherlands Environmental Assessment Agency (NEAA or MNP in Dutch), with special reference to the Nature Outlook and Environmental Outlook, each produced once in four years. Maps also form an integrated part of the annually produced Nature Balance and Environmental Balance. Maps and map comparisons play a role in the production of all these documents. Maps may be compared for a number of reasons, of which six will be named here: (i) the different economic/demographic scenarios forming the basis of maps, (ii) the detection of temporal changes, (iii) the calibration/validation of land-use models, (iv) hot-spot detection, (v) uncertainty analysis, and (vi) the different methodologies/models generating a set of maps.

This report addresses the problem of quantifying map similarities and dissimilarities. Our main focus is on maps denoted as 'categorical' or 'nominal'. Comparison is most complicated for the type of map responding to the question: how can we quantify similarities and dissimilarities in maps with nominal categories such as 'grassland', 'residential', 'infrastructural' or 'recreational'?

Here we will describe a number of five map-comparison techniques. These techniques differ in mathematical approach ('no math', 'cell by cell', two types of 'fuzzy', 'single-map statistics') and apply to different types of maps (nominal, ordinal, ratio and interval scales. Special attention is given to the comparison of maps based on *fuzzy-set calculation rules*. Here, both fuzziness of category and fuzziness of location are accounted for. The rationale is that fuzzy-set map comparison is very close to human judgement, as shown in an heuristic Internet experiment.

A variety of case studies show examples of the comparisons. Cases vary from land-use modelling for the year 2020 to historical reconstruction of worldwide agriculture from 1700 up to the present. Special topics are the role of map comparison in the calibration and validation of land-use models, and the search for consistency in land-use maps produced by Statistics Netherlands (CBS).

The MAP COMPARISON KIT (MCK) software plays a major role in the comparisons reported here. The MCK, a software package applied to 'state of the art' map comparison, contains all examples brought forward in the report. The software, developed by order of NEAA, was fully designed by the Research Institute for Knowledge Systems (RIKS). The software will be made publicly available on the RIVM website early 2004 ([www.rivm.nl](http://www.rivm.nl)). The MCK software can be concluded to be a powerful tool in map comparison applications. MCK is unique in having integrated four separate comparison techniques.



# 1. Introduction

by H. Visser, A.C.M. de Nijs and A. Hagen

## 1.1 The role of maps in environmental research

The use of (land-use) maps is important for a wide range of products issued by the Netherlands Environmental Assessment Agency (NEAA or MNP in Dutch), with special reference to the Nature Outlook and Environmental Outlook, each produced once in four years (see MNP (2002a) and MNP (2000), respectively). Other examples are found in MNP (2001) and MNP (2002c). Maps also play a role in the annual Environmental Balance and Nature Balance. Please see below for an example from Nature Outlook 2 , 2000 - 2030.

NEAA is required to report regularly on the current status and future trends in the environment and nature. Nature Outlook 2 presents a scenario analysis based on spatially detailed sketches of the Netherlands in 2030. These land-use maps were generated by a software tool called *the Environment Explorer* (De Nijs *et al.*, 2002). Nature Outlook 2 examines the effects of future developments in society on nature and landscape by means of four scenarios based on the so-called IPCC- SRES scenarios. The Nature Outlook 2 scenarios describe the development of Dutch society on the basis of two different contradictory trends: ‘globalization versus regionalization’ and ‘individualization versus co-operation’.

The four scenarios in Nature Outlook 2 have been converted to future land use using the present land use and the developments of the various land uses according the scenarios. The spatial quantification results in four maps showing the land use in 2030. In essence, the maps are sketches of the Netherlands in 2030, defined by the principles and assumptions of the scenarios, showing the results of current trends and potential transitions. The maps form the basis for subsequent emission, distribution and effect analysis in the Nature Outlook 2. The maps characterize the potential effects of the four scenarios on naturally and culturally valuable landscapes and reserves in the Netherlands. An example is given in **Figure 1**.

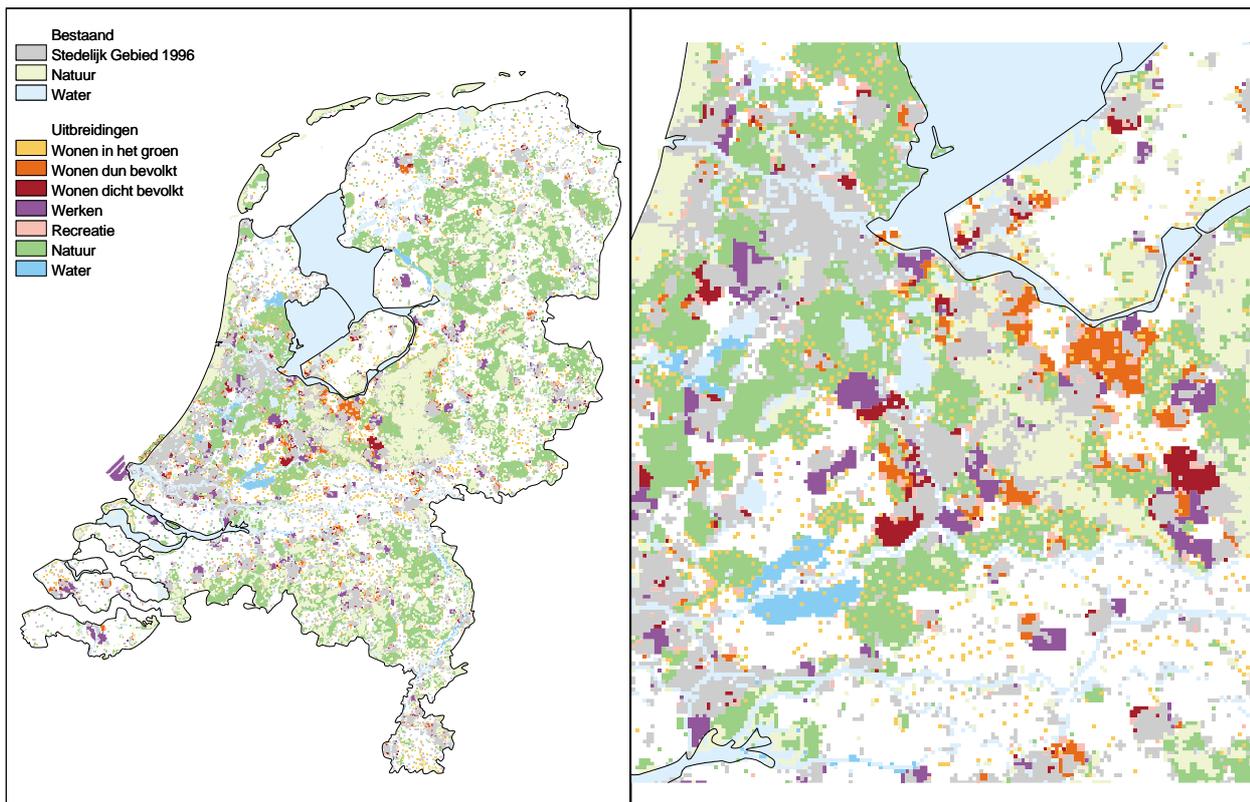


Figure 1 Land use in 2030 in the scenario 'Individual World' on a national scale (left panel) and regional scale (right panel) (Source: De Nijs *et al.*, 2002).

## 1.2 Why compare maps?

Maps may be compared for a number of reasons. We have chosen to describe six of these:

- Comparison of model output generated under different **scenarios** and assumptions.
- To detect temporal changes. Many maps have a temporal dimension. For example, in Chapter 5 we will show land-use maps from Statistics Netherlands (CBS) for 1989, 1993 and 1996. It would be interesting to study these land-use maps to find out where and for which categories major changes have occurred. Another example of temporal changes will be given in section 4.2.
- To calibrate/validate land-use models. Land-use models such as the Environment Explorer (De Nijs *et al.*, 2002) and the Land Use Scanner (Scholten *et al.*, 2001), generate land-use maps starting at an actual observed land-use map (here the CBS map for 1989). How well do these models predict future developments and how can we optimize model output to unknown parameters in the model? For such calibration problems we need an objective measure for map (dis)similarity. In fact, map comparison may be seen as giving a *Goodness-of-fit measure*. We will give an example of the procedure in section 5.1.

- Besides as a Goodness-of-fit measure map comparison has another role in the calibration/validation procedure. A map comparison method may give information about the **nature of the differences** between model map and reality map. Insight in these differences help in estimating better, adjusted, parameter values.
- To detect hot spots. With hotspot detection is meant the process of finding in a haystack of spatial data those areas that require further attention, say the **problem areas**. A map comparison method that is capable of distinguishing policy-relevant changes can be used for hotspot detection. For example, map comparison methods may be capable of recognizing changes that are related to land degradation.
- To perform uncertainty analyses. There are many sources of error in maps. By comparing model output to a reference map, such errors may be detected and quantified. In fact, calibration is a form of uncertainty analysis. The study of changes in category definitions in land-use maps is also an example (section 5.2).
- To compare different methodologies. Maps may be generated by different models. How well do these maps compare and if they differ, where are the differences located and how can we quantify them? An example of such maps is given in section 4.2, where historical maps for agriculture are studied. Two independent research groups generated maps for the years 1700, 1850 and 1990.

Here we will introduce a number of five map-comparison techniques. These techniques differ in mathematical approach (no math, 'cell by cell', 'fuzzy' or single-map statistics), and apply to different types of maps. Our main focus will be on maps denoted as 'categorical' or 'nominal'. Comparison is most complicated for these types of maps: how can we compare 'grassland' with 'residential', or 'glasshouses' with 'recreational'? Still, we are looking for exact quantification of differences between maps containing such categories.



*Land-use models such as the Environment Explorer and the Land Use Scanner predict land-use changes given scenarios for future (economic) developments. Changes as given above in the four cyclorama pictures should be predicted as much as possible. Map comparison may help to locate and quantify differences in model output and actual observed changes. Quantification of changes is important for the calibration and validation of land-use models.*

*The pictures are taken from exactly one point near the village of Zetten:*

- 1. April 1999 (upper panel). Shown is an orchard located in the Betuwe region. Here, a new railway line for freight transport (the so-called Betuwe line) is being built.*
- 2. April 2000 (second panel). Trees have been cut down for a temporary roadway (the A15).*
- 3. May 2001 (third panel). Groundwork for the detour of the A15 roadway.*
- 4. November 2001 (lower panel). Asphalt has been laid. The Betuwe railway line will eventually pass through this area.*

*Photos: S. Swart.*

### 1.3 From cell-by-cell to fuzzy

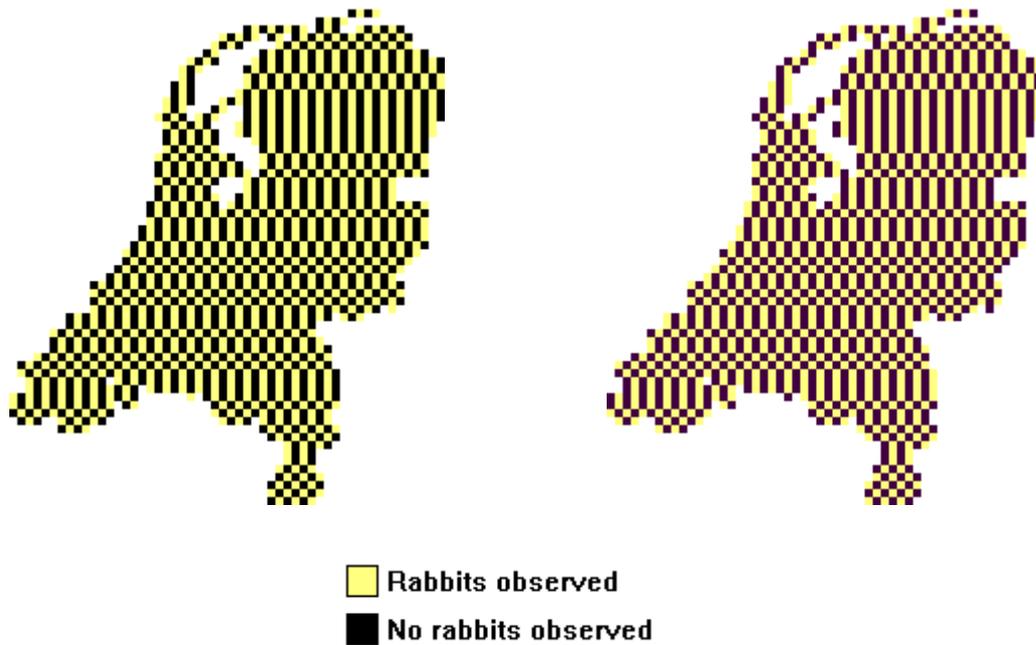
Growth of high-resolution spatial modelling, geographical information systems and remote sensing has increased the need for map comparison methods. Good comparison methods will be needed to perform calibration and validation of spatial results in a structured and controllable manner. The importance of map comparison methods has been recognized and has stimulated growing interest among researchers (Monserud and Leemans, 1992; Metternicht, 1999; Winter, 2000; Pontius, 2000; Pontius and Schneider, 2001; Power, Simms and White, 2001; Hagen, 2003).

For most purposes, visual, human comparison still outperforms automated procedures. When comparing maps the human observer takes many aspects into consideration without deliberately trying. Local similarities, but also global similarities, logical coherence and patterns are recognized. Map-comparison methods performed by software usually capture one of these aspects but overlook the others. Furthermore, they generally lack the flexibility to switch from one aspect to the other when the data requires so.

The best example of this rigidity is the cell-by-cell comparison of two checkerboards. The first board has a white field in the upper left corner, the second a black field. The average observer would immediately recognize the two boards as being highly similar in quality. However, a cell-by-cell comparison method would find a black cell where a white one is expected and vice versa. Hence, total disagreement would be concluded. A checkerboard example is given in **Figure 2**.

Despite these clear disadvantages, there are situations where automated map comparison is preferred above visual map comparison, for example, where an automated procedure can save time and human effort. A more important reason for choosing automated comparison is that automated procedures are explicitly defined and therefore objective and repeatable. Thus the method can be analysed and evaluated, and the results verified. A visual map comparison will always be subjective and often intuitive. The outcome of a visual map comparison can therefore depend on the person performing the comparison.

The main map-comparison method presented in this report was primarily developed for use in the calibration and validation process of cellular models for land-use dynamics. The method is based on the **fuzzy-set calculation rules** (Zadeh 1965). Several authors have addressed the potential of fuzzy-set theory for geographical applications (Cheng, Molenaar and Lin, 2001; Fisher, 2000). In the past fuzzy-set theory was used to assess the accuracy of map representations and for map comparisons (Metternicht, 1999; Power, Simms and White, 2001).



*Figure 2* Two maps showing the observed presence of rabbits in the Netherlands. Grid size is 5 x 5 km.

For some reason the presence of rabbits has been shifted exactly one cell between both maps. A cell-by-cell comparison method would qualify both maps as *totally different*. However, an average observer would qualify both maps as *highly equivalent*. In this sense, map comparison based on *fuzzy-set calculation rules* is very similar to human judgement.

The objective of fuzzy-based map comparison is to find a method that to some extent mimics human comparison and gives a detailed assessment of similarity. The method introduced in this report is aimed at comparing nominal and ordinal raster maps. The assessment results are spatial and gradual. Additionally, an overall figure for similarity, the *Fuzzy Kappa statistic*, is aggregated from the detailed spatial results.

Although fuzzy-based map comparison has clear advantages over the more traditional cell-by-cell methods, as we will point out later, the latter still yields results showing insight into many cases. For this reason we will follow *two tracks* in this report. The strong points of both cell-by-cell comparison techniques and fuzzy-based map comparison will be combined to get 'the best of two worlds'.

## 1.4 The MAP COMPARISON KIT

A central role in this report is played by the MAP COMPARISON KIT (MCK) software. The MCK is a software package for 'state of the art' map comparison developed by the Research Institute for Knowledge Systems (RIKS) (Hagen, 2002a) by order of NEAA. The MCK user's guide was provided by RIKS (2003a).

This report can be seen as a practical guide to the application of map comparison using the MCK software. Five map comparison techniques are described, four of which are implemented in the software (sections 2.2 through 2.5). All examples in this report, except for the one found in Appendix A, are calculated by the MCK.

MCK software was developed in the first place for environmental applications within NEAA. All case studies in sections 4.2 to 4.6 are taken from NEAA's daily practice. However, the software is employable for many GIS users outside NEAA as well. Therefore, the software will be made available to the general public at the beginning of 2004, via the RIVM website.

## 1.5 This report

As stated in section 1.2, this report is directed to map-comparison techniques based on different mathematical approaches as implemented in the MCK software. These techniques vary with the type of maps used, from nominal and ordinal scales to interval or ratio scales. Chapter 2 presents an overview of five techniques. Theoretical backgrounds are shortly described for all cases. More extensive descriptions can be found in the literature cited. Methods in sections 2.2 to 2.5 are implemented in the MCK. The single-map approach in section 2.6 (and continued in Appendix A) can be followed using such software packages as S-PLUS SpatialStats and INFLUENCE.

In Chapter 3 we give a short overview of the MAP COMPARISON KIT software (section 3.1), and two kinds of software implementations, additional to the MAP COMPARISON KIT (section 3.2). Section 3.3 will provide a description of software available for single-map statistics.

Chapter 4 presents a number of practical examples and NEAA case studies. It starts off with a humorous illustration of the methods described in sections 2.3 and 2.4, followed by NEAA case studies on handling nominal maps in sections 4.2 and 4.3, ordinal maps in section 4.4 and ratio maps in sections 4.5 and 4.6. We proceed in Chapter 5 to describe two specialized applications: calibration of land-use models (section 5.1), and identification and correction of inconsistencies in land-use maps (section 5.2). The report ends with conclusions and ideas for further research.

In Appendix A we describe an example of the single-map-statistics approach as described in section 2.6. The rationale behind enrichment factors is given, along with an example. Appendix B gives a detailed description of software additional to the MAP COMPARISON KIT. As an extension to Chapter 2, Appendix C states a case for the Fuzzy-set map comparison approach showing great similarities to human judgement in comparing maps. This strengthens the argument for applying methods based on fuzzy principles rather than cell-by-cell statistics.

## 2. Map comparison - methods

We can distinguish five methods for comparing maps:

1. visual map comparison
2. cell-by-cell map comparison
3. map comparison based on Fuzzy-set calculation rules
4. map comparison based on Hierarchical fuzzy pattern matching
5. map comparison based on single-map statistics

The first method, visual inspection of maps, is well-known, and may seem almost trivial. However, experience has taught us that careful inspection of data may reveal many interesting or 'suspect' characteristics of the maps. These characteristics could be overlooked if one applies advanced (statistical) techniques directly to the spatial data (see section 2.2). The second method, cell-by-cell with Kappa statistics, is also well-known. Kappa statistics are widely used in land-use modelling nowadays. We will extend the traditional Kappa statistic by two advanced statistics: the 'Kappa location', designed by Gill Pontius (2000), and the 'Kappa histo', designed by Alex Hagen (2003) (see section 2.3).

The third and fourth methods introduce a relatively new vision on map comparison techniques that are especially suited for nominal maps. Both methods are based on fuzzy-set theory.



*Advertisement for a laser eye correction: left panel before operation, right panel after operation. In fuzzy map comparison we perform exactly the opposite 'operation' to the one above, i.e. we blur the images before making a comparison.*

Here, map comparison is no longer solely based on cell-by-cell, but also on the neighborhoods of corresponding cells in two maps. In fact, the borders of categories are made vague (or fuzzy) before map statistics are calculated (see sections 2.4 and 2.5).

The fifth method, comparison on the basis of single-map statistics, is simple, yet seldom applied in practice. Any statistic that can be calculated for one single map can be used to compare this characteristic for a set of two or more maps. Interesting examples are expressions for the fragmentation of nature areas, the estimation of trends in a particular category of interest and the application of neighborhood characteristics for the calibration of cellular automata models (section 2.6). We will start with a few definitions.

## 2.1 Definition of measurement scales

There are a number of different types of maps, depending on the scale of measurement. In this report we use the terminology as summarized on the **next page**. Maps are distinguished according to the following scales: nominal, ordinal, interval and ratio.

These basic measurement scales may also appear in mixed forms, especially relevant for nominal and ordinal scales. For example, the Environment Explorer presents 16 land-use categories; of these, 14 are on a nominal scale and 2 on an ordinal scale: ‘sparse residential’ and ‘dense residential’.

To be complete here, we should mention that some maps are *hierarchically* classified, with classes and subclasses. A class may be ‘agriculture’, while subclasses may be ‘glass houses’, ‘arable farming’ and ‘cattle breeding’. A data set such as *CORINE* employs a class system with four hierarchical levels.

## Measurement scales

Throughout this report we use four types of measurement scales, as defined below (adapted from: [www.douglas.bc.ca/psychd/handouts/measurement\\_scales](http://www.douglas.bc.ca/psychd/handouts/measurement_scales)).

### *Nominal scale*

This is the simplest and most elementary type of measurement. The purpose of a nominal scale is to differentiate one object from another. Some examples of nominal data include the categories of male/female or different religions. Another way to think of a nominal scale is that it is simply a classification system. For instance, we can define 16 land-use categories in a country map. With nominal data we cannot compute averages as these data lack the necessary properties to do so.

### *Ordinal scale*

An ordinal scale means that the measurements now contain the property of order. We can not only specify the difference between one object and another, but also the *direction* of difference. This enables us to make statements using the phrases 'more than' or 'less than'. Ranking is an example of an ordinal scale. It simply allows us to specify the direction of the difference, not the quantity of the difference.

### *Interval scale*

Interval scales have the same characteristics as nominal and ordinal scales, i.e. the ability to classify and to indicate the direction of the difference. However, this scale has the added benefit of equality of units. These equal distances between observation points permit us to make statements about the direction of the differences and to indicate the amount of the difference. The Fahrenheit temperature scale is an example of an interval scale because of its equality of units. Interval scales typically occur in psychological or social studies.

### *Ratio scale*

The ratio scale incorporates all the characteristics of the interval scale with one important addition - an absolute zero. Using an absolute zero point, one can make statements involving ratios of two observations such as 'twice as long' or 'half as fast'. For example, the Kelvin temperature scale has a ratio scale because a temperature of 0 Kelvin indicates the absolute zero point of temperature. Typically, ratio scales apply to physical measurements. Examples for maps are chances between 0.0 and 1.0, as generated by RIVM's 'Natuurplanner' (Nature Planner), or a percentage of grids with coverage of a certain category, as in the RIVM's 'Ruimtescanner' (Land Use Scanner).

## 2.2 Visual map comparison

There is an old adagium in data analysis: always look at your data carefully, before doing advanced analyses. For map comparison, we have to visualize our set of maps and look at differences. Of course, this approach is subjective but it serves as a good start for any analysis. By zooming in we can detect interesting, as well as ‘suspect’ map differences.

The following description is the result of a visual comparison of the two fictitious maps given in **Figure 3**.

*‘I see two maps that are not identical; however they clearly depict the same area. Both maps have a river network that is roughly the same but show minor differences all over. Maybe the maps are of the same area but at different times and the river has changed its course a bit. There is one big difference in the river network; this is found in the upperleft part of the map. Here the second map has a river branch that in the first map is missing.*

*I recognize two urban clusters on both maps. They are located at the same location in both maps. The orientation of the upper cluster seems to be different in the two maps. The clusters, maybe towns or villages, have approximately the same size in both maps.*

*Also in both maps there is a speckled pattern of city, park and open area. I wonder if this is a bungalow park or something else. The precise location of park and city is completely different in both maps and the total area taken in by this cluster is larger in the second but denser in the first map.’*

Other opinions on Figure 3 could be:

*'These maps are similar but somehow the second one seems to be more artificial.'*

*'The difference between these two maps are highly significant, I suppose they are of the same area but they must be at least ten years apart in time.'*

*'These maps are practically identical, I can not imagine them to be depicting two different areas.'*

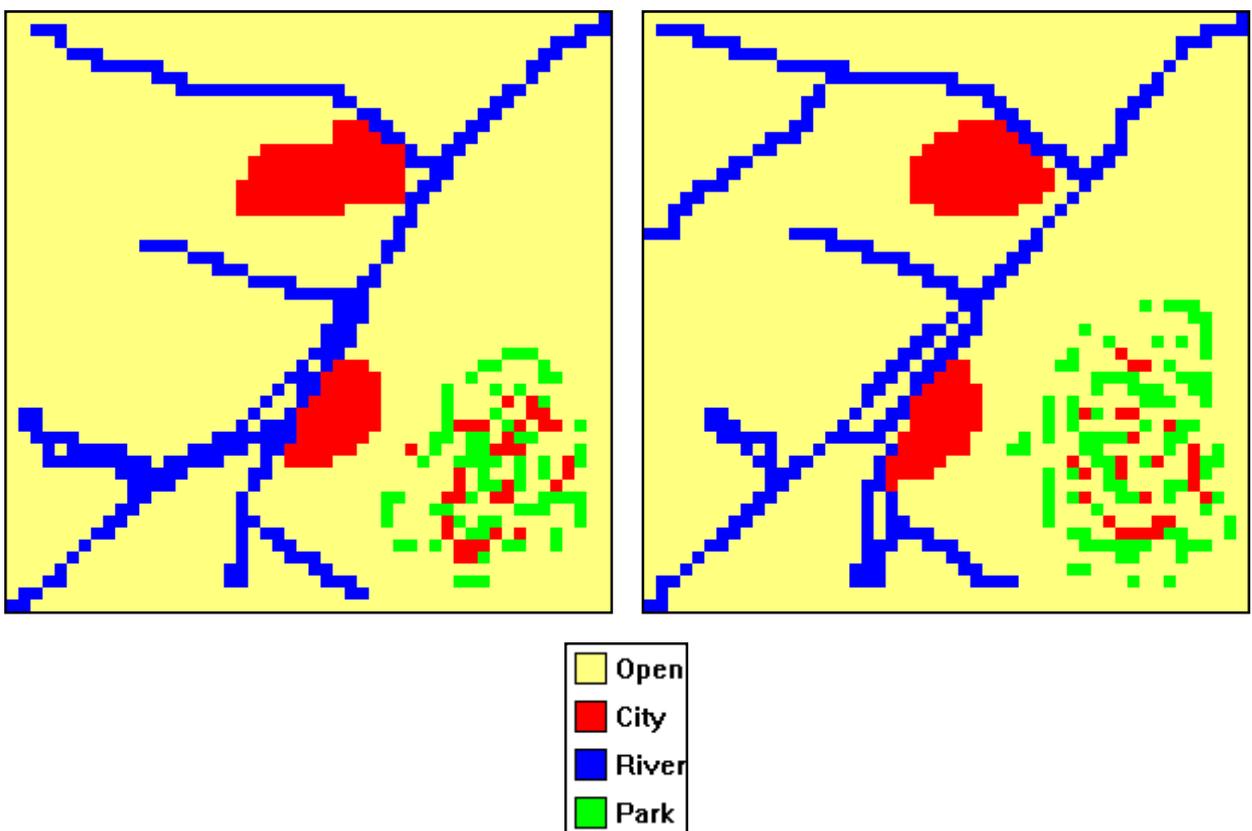
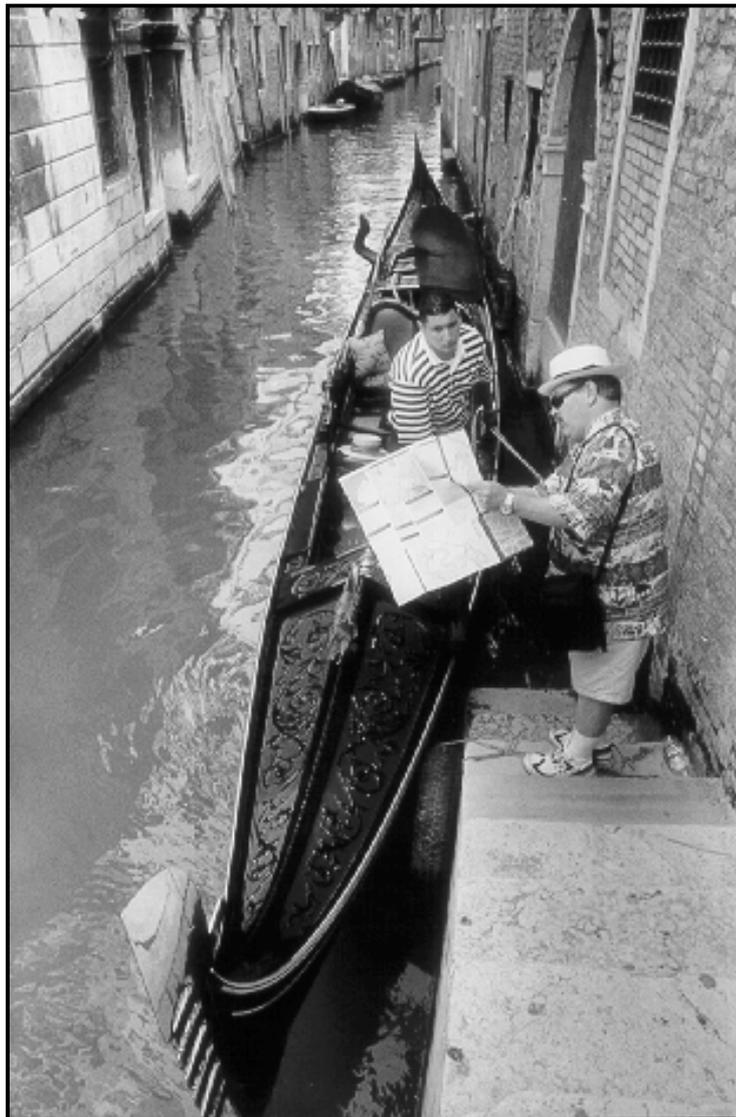


Figure 3 Two fictitious land-use maps.



*Visual map comparison has strong and weak sides. Intuition and personal interpretation clearly play a subjective role. However, visual inspection of maps has a logic and integrality which one misses in standard automated map-comparison techniques.*

*Fuzzy-set map comparison as introduced in section 2.4, tries to combine the strong sides of visual map comparison, and the objectivity and repeatability of an automated procedure. Please see Appendix C for an Internet map-comparison test.*

*Photo: H. Visser.*

## 2.3 Comparing maps cell by cell - Kappa statistics

by A.Hagen

### 2.3.1 Equal or unequal, that's the question

For maps on interval or rational scales, differences can be expressed numerically on a cell-by-cell basis. However, for maps on nominal or ordinal scales, we only have the outcome 'equal' or 'unequal' for identical grid points in both maps. Therefore it would be helpful to project a difference map between two maps with these binary outcomes for all grid cells available. Thus, the difference maps shows for all grid pairs in both maps the outcome 'equal' or 'unequal'.

We can also construct a series of difference maps *per category*. Focusing on one particular category, we can identify four outcomes: 'present in none of the maps', 'present in both maps', 'present in map 1, not in map 2' and 'present in map 2, not in map 1'.

A fictitious example is given in **Figure 4**. The upper-left panel shows 'map 1', the upper-right panel 'map 2'. The lower-left panel shows a simple equal-unequal result map. For the lower-right panel we focused on the category 'River'. The other three categories are aggregated to a category 'other'. The white area on the result map means that the category 'other' is identical in both maps, i.e. is not 'River' in both maps.

### 2.3.2 Advanced use of Kappa statistics

The Kappa statistic is often used to assess the similarities between observed and predicted results. It is not only applied to geographical problems (e.g., Pontius, 2000; Monserud and Leemans, 1992) but to many other fields such as medical and social sciences. As a result much has been published about the Kappa statistic and its functionality extensively discussed (Carletta, 1996; Fielding and Bell, 1996; Lantz and Nebenzahl, 1996; Maxwell, 1977).

This section will shortly discuss the Kappa statistic and the contingency table forming its basis, followed by the introduction of a few related statistics, and suggestions for practical use of Kappa statistics. The description is taken from Hagen (2002b).

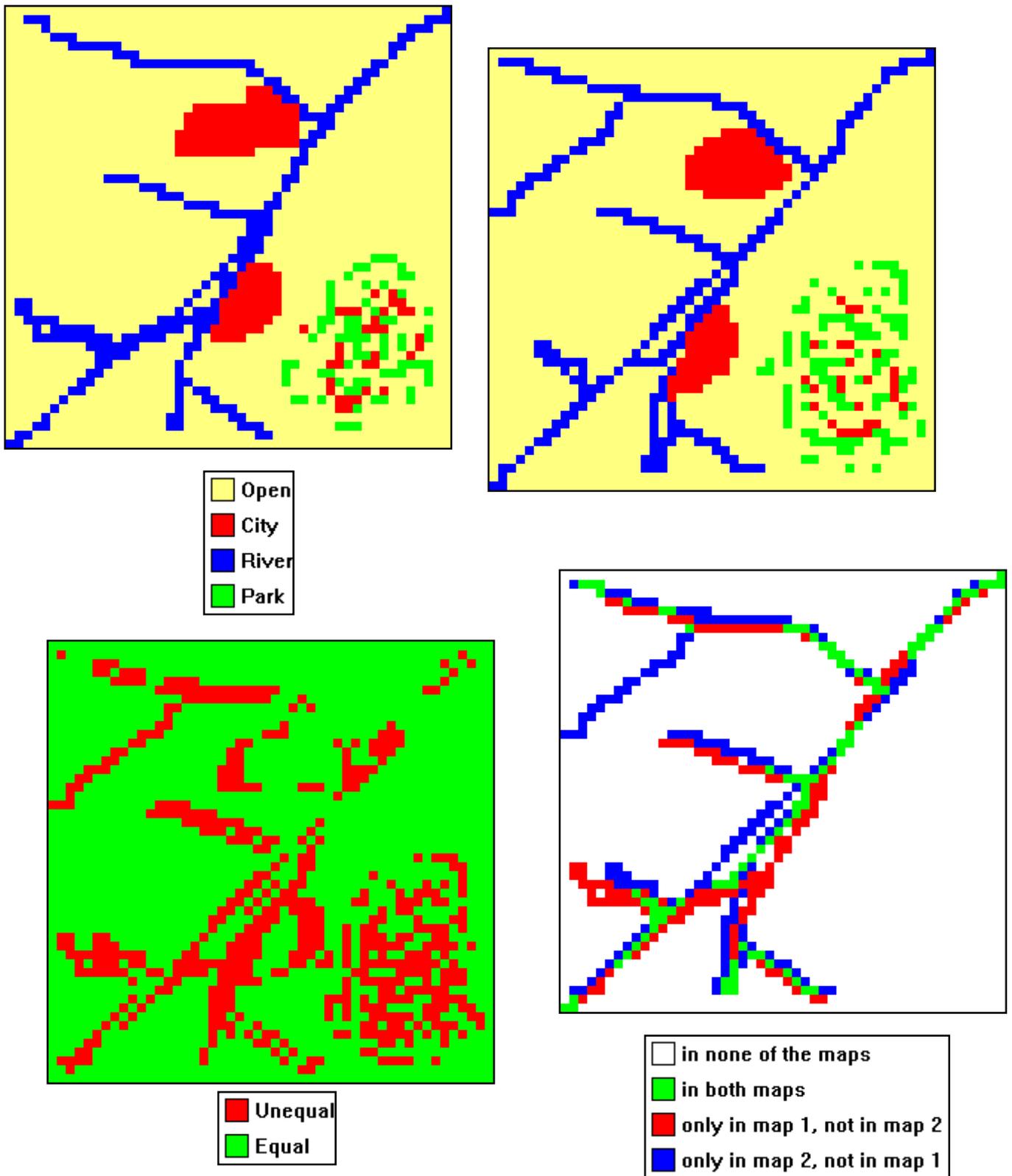


Figure 4 Upper two panels show two fictitious land-use maps (identical to Figure 3). The lower left panel shows a cell-by-cell comparison, yielding equal (green) or unequal (red). The lower right panel shows a cell-by-cell comparison for the category 'river'. There are four possible outcomes for each cell, as shown in the lower right legenda.

### 2.3.2.1 Contingency table

The calculation of Kappa is based on the so-called contingency table (sometimes also referred to as confusion matrix). **Table 1** shows the generic form of a contingency table. The table details how the distribution of categories in map A relates to that of map B. The cells contain a value representing the fraction of the cells in the map: in map A by the category specified in the matrix **row**, and in map B by the category specified in the matrix **column**. For example, a value of 0.10 for  $p_{12}$  indicates that 10% of the total mapped area, belonging to category 1 in map A, is found in category 2 of map B.

The last row and column give the column and row totals. Each **row** total represents the total fraction of cells of the related category in map A. Similarly, each **column** total represents the total fraction of cells of the related category in map B. All fractions taken collectively make up the whole map, with the total sum equal to 1.

*Table 1 The contingency table in its generic form (Monserud and Leemans, 1992).*

		Map B categories				Total
		1	2	...	c	
Map A categories	1	$p_{11}$	$p_{12}$	...	$p_{1c}$	$p_{1T}$
	2	$p_{21}$	$p_{22}$	...	$p_{2c}$	$p_{2T}$
	c	$p_{c1}$	$p_{c2}$	...	$p_{cc}$	$p_{cT}$
Total		$p_{T1}$	$p_{T2}$	...	$p_{Tc}$	1

A number of statistics can be derived on the basis of the contingency table. The following three statistics are applied here:

1. P(A) stands for Fraction of Agreement and is calculated according to Equation (1).
2. P(E) stands for Expected Fraction of Agreement subject to the observed distribution and is calculated according to Equation (2).
3. P(max) stands for Maximum Fraction of Agreement subject to the observed distribution and is calculated according to Equation (3).

$$P(A) = \sum_{i=1}^c p_{ii} \quad (1)$$

$$P(E) = \sum_{i=1}^c p_{iT} * p_{Ti} \quad (2)$$

$$P(\max) = \sum_{i=1}^c \min(p_{iT}, p_{Ti}) \quad (3)$$

### 2.3.2.2 Kappa statistics

In many situations it is preferential to express the level of agreement in a single number. The Kappa statistic may be a suitable approach if the comparison consists of a number of pairwise comparisons (Carletta, 1996). The essence of the Kappa statistic is that the fraction of agreement  $P(A)$  is corrected for the fraction of agreement statistically expected from randomly relocating all cells in the map. Thus, this expected agreement is based on random location of subject to the observed distribution; this is referred to as  $P(E)$ .

The Kappa statistic is defined according to Equation (4):

$$K = \frac{P(A) - P(E)}{1 - P(E)} \quad (4)$$

Note that Kappa may become negative if the fraction of agreement  $P(A)$  is worse than random relocation of all cells  $P(E)$ .

### 2.3.3 Kappa dissected into K-histo and K-location

Pontius (2000) clarified the Kappa statistic as confounding similarity of quantity with similarity of location. In this sense ‘quantity’ means the total presence as a fraction of all cells in a category taken over the whole map. ‘Location’ means the spatial allocation of quantity over the map. Pontius introduces two statistics to separately consider similarity of location and similarity of quantity. The statistic for similarity of quantity is called **K-quantity**, but the application of this statistic leads to many practical problems. However, the statistic for similarity of location is very informative because it gives the similarity scaled to the maximum similarity which can be reached with the given quantities.

**K-location** is calculated according to Equation (5):

$$K_{\text{location}} = \frac{P(A) - P(E)}{P(\text{max}) - P(E)} \quad (5)$$

An alternative expression for the similarity of the quantitative model results is the maximal similarity that can be found on the basis of the total number of cells taken in by each category. This is called P(max). P(max) can be put in the context of Kappa and K-location by scaling it to P(E). The resulting statistic, newly introduced here, is called **K-histo** since it is a statistic that can be calculated directly from the histograms of two maps.

K-histo is defined by Equation (6):

$$K_{\text{histo}} = \frac{P(\text{max}) - P(E)}{1 - P(E)} \quad (6)$$

The powerful property of the K-histo definition is that Kappa is now defined as the product of two factors:

$$K = K_{\text{histo}} * K_{\text{location}} \quad (7)$$

The first factor is K-location, a measure for the similarity of spatial allocation of categories for the two maps compared. The second factor is K-histo, which is a measure of the quantitative similarity of the two maps.

Besides calculating Kappa statistics for all categories combined, Kappa statistics per category can be calculated as an option. For a categorical Kappa statistic, the two maps are transformed to a map consisting of only two categories. The first new category is the category for which the individual Kappa statistic is derived; the second is the combination of all other categories.

### ***2.3.3.1 Relative Kappa statistics***

A typical map comparison problem is the question how well a map generated by a model (the Model Map) resembles a real map (the Reality Map). The Kappa statistic can be of use here. By itself, however, it offers insufficient information: a Kappa statistic with a value 0.7 may be considered very high in one case but can indicate a poor result in another. For an indication of how well two maps look alike, a reference level for similarity is needed. This reference level can be obtained from a Reference Map, for instance, in the form of a historical map.

The procedure is as follows. In first instance the Model Map is compared to the Reality Map. This comparison yields several statistics: Kappa, K-histo and K-location. The same operation is performed for the Reality Map and a Reference Map. This comparison also yields values for Kappa, K-histo and K-location. Finally, the individual comparison results are combined, and the similarity between the Model Map and the Reality Map can be expressed *relative* to the similarity of the reference map (cf. Figure 9 in section 2.4.5).

### 2.3.4 Example

The multi-method similarity assessment is applied to a validation case of model results. The particular model is a constrained cellular automata applied to the study of the urban development of Dublin as part of the Murbandy project (White *et al.*, 2000). The objective of the case was to compare model results with observed data. The two maps are displayed in **Figure 5**.

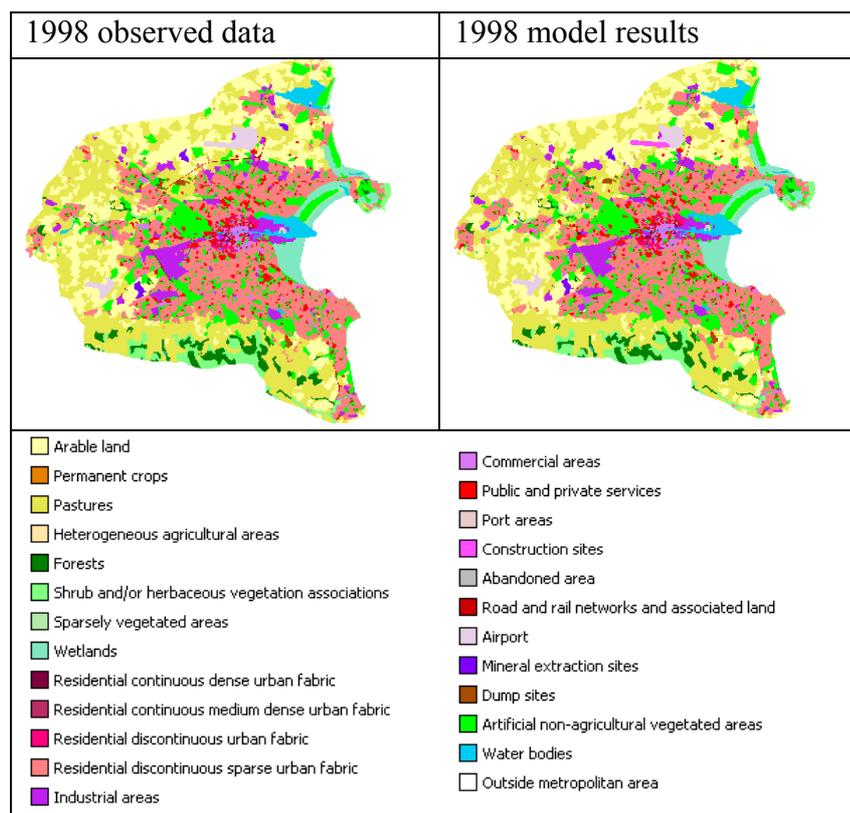


Figure 5 Observed and simulated maps of the metropolitan area of Dublin in 1998.

With the Kappa-related statistics it is possible to recognize the contribution per individual category and also to distinguish between both similarities due to quantity and those due to location. The result of this analysis can be found in **Table 2**.

The results presented in the table suggest that although a little improvement can still be made, most of it can be expected from improving the spatial allocation. The categories with relatively weak spatial allocation are 'Residential continuous dense urban fabric' and 'Construction sites'. The relatively low scores for 'Road' and 'Airport' are in accordance with the observations made on the comparison map.

*Table 2 Detailed Kappa results, overall and per individual category.*

	<b>Kappa</b>	<b>K-location</b>	<b>K-histo</b>
<b>Overall</b>	0.96	0.97	0.99
Arable land	0.95	0.96	0.98
Pastures	0.94	0.96	0.98
Forests	1.00	1.00	1.00
Shrubs and/or herbaceous vegetation associations	1.00	1.00	1.00
Sparsely vegetated areas	1.00	1.00	1.00
Wetlands	1.00	1.00	1.00
Residential continuous dense urban fabric	0.78	0.78	1.00
Residential continuous medium-dense Urban fabric	0.95	0.95	1.00
Residential discontinuous urban fabric	1.00	1.00	1.00
Residential discontinuous sparse urban fabric	0.91	0.91	1.00
Industrial areas	0.96	0.96	1.00
Commercial areas	0.86	0.86	1.00
Public and private services	0.95	0.95	1.00
Port areas	0.85	0.85	1.00
Construction sites	0.00	0.00	0.08
Road and rail networks, and associated land	0.43	0.82	0.53
Airport	0.88	1.00	0.88
Mineral extraction sites	0.97	1.00	0.97
Dump sites	0.99	0.99	1.00
Artificial non-agricultural vegetated areas	0.93	0.93	1.00
Water bodies	1.00	1.00	1.00
Outside metropolitan area	1.00	1.00	1.00

## 2.4 Fuzzy-set map comparison

by A. Hagen

In this paragraph fuzzy-set theory, as introduced by Zadeh (1965), will be applied to compare nominal maps. In order to consider fuzziness in the maps it is necessary to change the way in which cells are represented. Instead of one single category or value per cell, each cell is characterized by a membership vector. Each element in the vector states the extent of membership per category using values between 0 and 1,.

Two sources of fuzziness are considered, the first is fuzziness due to vague distinctions between categories, and the second is fuzziness due to a gliding scale of severity of spatial error. The comparison method is documented more extensively in Hagen (2002b).

### 2.4.1 Considering categorical similarity

Many maps show vagueness in the definition of categories. This is especially true if some or all categories on the map have, in fact, an ordinal definition, such as the categories 'high', 'medium'-and 'low'-density residential area on a land-use map. It might often be that boundaries between such categories are less clear-cut than what seems to be the case from the legend. This fuzziness can be made explicit in the vector describing the cell, by giving elements that correspond to similar categories and higher membership values. **Table 3** gives an example how the fuzziness of the categories can be expressed in the membership vector.

*Table 3 An example of fuzzy representation of ordinal data.*

Category	Nr.	Category vector
High density residential	1	( 1 0.4 .2 0 0 )
Medium density residential	2	( 0.4 1 .4 0 0 )
Low density residential	3	( 0.2 0.4 1 0 0 )
Agriculture	4	( 0 0 0 1 0 )
Industry	5	( 0 0 0 0 1 )

## 2.4.2 Considering proximity of similar cells

Proximity of similar cells can also be expressed in the membership vector. Cells within a certain distance (the neighborhood) of a central cell influence the fuzzy representation of that cell. To achieve this, the proximity of categories is considered to contribute to the degree of membership of those categories. The different membership contributions of the neighbouring cells are combined by calculating the union according to fuzzy-set theory. This is expressed in Equation (8) for a map with N categories and considering a neighborhood consisting of C cells.  $m_i$  stands for the value of the membership function at the i-th cell in the neighborhood and is calculated according to a distance decay function.

$$V_{\text{fuzzy}} = \begin{pmatrix} F_1 = \text{Max}(\mu_{1,1} * m_1, \mu_{1,2} * m_2, \dots, \mu_{1,C} * m_C) \\ F_2 = \text{Max}(\mu_{2,1} * m_1, \mu_{2,2} * m_2, \dots, \mu_{2,C} * m_C) \\ \vdots \\ F_C = \text{Max}(\mu_{N,1} * m_1, \mu_{N,2} * m_2, \dots, \mu_{N,C} * m_C) \end{pmatrix} \quad (8)$$

## 2.4.3 Comparison of fuzzy cells

The maps of fuzzy membership vectors obtained by considering proximity and categorical similarity are compared. The comparison algorithm is designed to evaluate similarity in accordance with human ‘intuitive’ criteria. This can be achieved by performing a two-way comparison, proceeding as follows: first, the fuzzy vector of cell A is compared to the category vector of cell B according to fuzzy-set theory. Next the category vector of cell A is compared to the fuzzy vector of cell B. Finally, *the lower of the two comparison results* establishes the similarity.

A similarity map is generated by applying the comparison cell-by-cell to the whole area. In this similarity map each cell has a value between 0 (for total disagreement) and 1 (for identical cells). **Figure 6** shows six situations that clarify this point; it should be noted that the exact value for the intermediate similarities (between total disagreement and identical) will depend on the membership function applied. The similarity values in the figure are based on a membership function of exponential decay with a halving distance of  $\sqrt{2}$  cells.

		<p>Situation 1: The value for similarity in the central cell will be low since the two cells (black and white) differ; furthermore, there are no cells of the same category in the neighbourhood.</p>
		<p>Situation 2: The value for similarity in the central cell will be intermediate since the two cells (black and grey) differ, however, there are cells of the same categories in the neighbourhood.</p>
		<p>Situation 3: As in Situation 2, the value for similarity in the central cell will be intermediate. The similarity will be less than in Situation 2 since the matching cells are found within a greater radius.</p>
		<p>Situation 4: The value for similarity of the central cell is equal to the one in Situation 3 since the matching cells are found within the same radius. The white cells do not influence the comparison.</p>
		<p>Situation 5: The value for similarity in the central cell will be low since the two cells (black and grey) differ; furthermore, there are no cells of the same categories in the neighbourhood.</p>
		<p>Situation 6: The value for similarity in the central cell will be high since the two cells match (both black), regardless of the circumstance that the neighborhoods (grey and white) are dissimilar.</p>
		<p>The calculated similarity of the central cell is 1</p>

Figure 6 Six situations comparing left and right maps, taking fuzziness of location into consideration.

In all six cases the similarity is given with respect to the central cell of the left map. Similarities with respect to the central cell of the right map are always equal or larger.

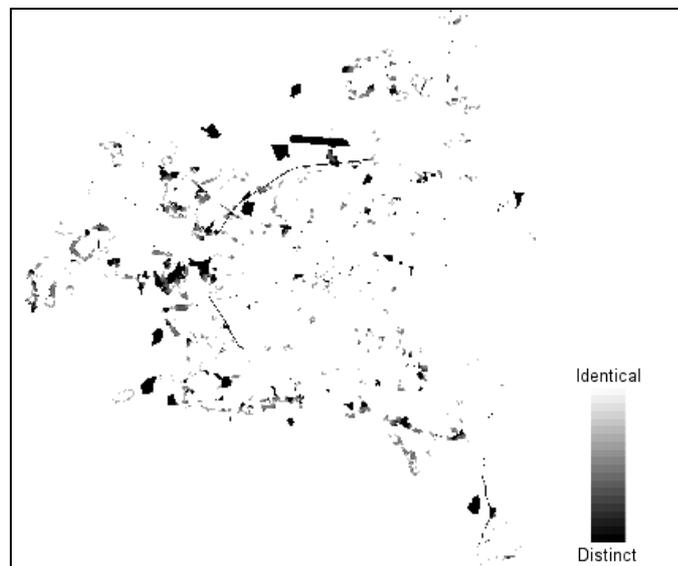
#### 2.4.4 Aggregated map results and overall similarity with Fuzzy Kappa

It is possible to aggregate the similarity map that results from the Fuzzy two-way comparison to an overall value of map similarity, for instance, by integrating the similarity values over the whole map. Subsequent division by the total area yields a result between 1 (for identical maps) and -1 (for total disagreement).

The outcome of the Fuzzy-set map comparison depends partly on the number of categories present and also on the numerical distribution of cells over those categories. In order to make the results of maps with different numerical distribution more comparable, we have introduced a statistic that corrects the percentage of agreement for the expected percentage of agreement. The statistic is similar to Kappa and is therefore called **Fuzzy Kappa**. For details please see Hagen (2003).

#### 2.4.5 Dublin example continued

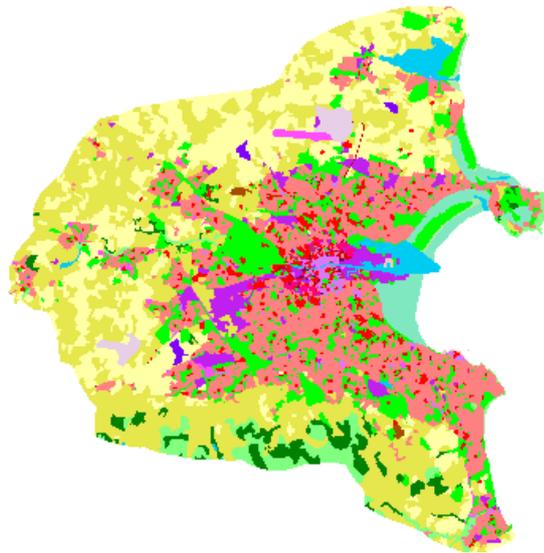
Here, we continue the Dublin example from section 2.3, having applied Fuzzy-set map comparison to the maps shown in Figure 5. **Figure 7** shows the results in the form of a comparison map indicating the level of agreement per cell. The membership function that was applied is one of exponential decay with a halving distance of two cells.



*Figure 7* Spatial assessment of similarity in the fuzzy set approach. The darker the area the more intensive the disagreement.

The result map can be used as an aid in to find the cause of the disagreement. For instance, we see large areas of strong disagreement in the north of the city where the model situates 'Commercial areas' where an 'Airport' is expected. The result map also clearly points out the 'Road and rail networks, and associated land', representing a motorway existing in reality but not foreseen by the model (the curved linear shape starting just south of the airport).

A reference level, sought for validation, was found in the map of observed data of 1988 (see **Figure 8**). This map was also used to show the initial situation in the simulation. If the 1998 model simulation is more similar to the 1998 observed data than the 1988 reference level does to the 1998 observed data, the validation is positive. Taking into account that changes in land use are only slight over a period of ten years, this is a strict validation.



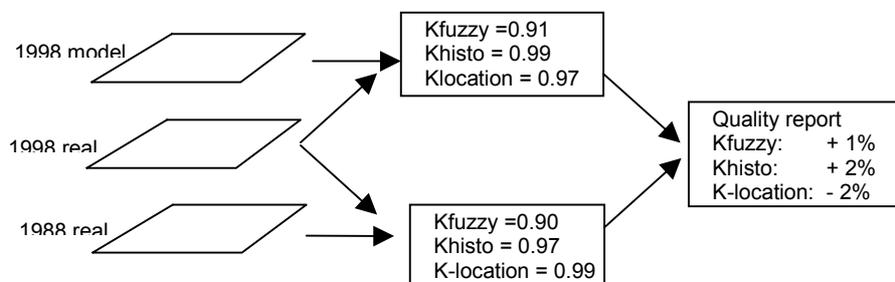
*Figure 8 The observed map of 1988, which functioned both as the Reference Map and a map showing the initial situation in the simulation.*



*Children playing in a residential discontinuous sparse urban fabric area of Dublin (category definition given in Table 4).*

*Photo: H. Visser.*

The comparison performed conformed to the method presented in section 2.3.3.1, with results presented schematically in **Figure 9**. The model was concluded as delivering a positive validation of results.



*Figure 9 Relative results of the comparison.*

## 2.5 Comparing maps by Hierarchical fuzzy pattern matching

The evaluation of spatial similarities and (land-use) change between two raster maps is traditionally based on cell-by-cell comparison techniques, as previously described. Along with the method based on fuzzy-set calculation rules, given in the preceding section, there are also other fuzzy methods in existence. One of these is the *Hierarchical fuzzy pattern matching*. This approach is based on the comparison of overlapping polygons (= groups of concatenated cells belonging to the same category). Polygons are then compared on the basis of a number of properties, such as size and overlap. These comparisons are based on a so-called *Fuzzy Inference System* (FIS).

The maps are compared at both local and global levels. This combination provides a robust alternative approach. Local matching determines the degree of containment of each unique polygon in the template map in terms of so-called fuzzy area intersections. Formally, the local agreement values are based on polygon property containments; calculations originate in a fuzzy logical Max-Min compositional algorithm.

A global agreement value can be derived by the fuzzy summation of the local matchings. This global agreement, compared to the Fuzzy-Kappa statistic, is denoted as *Global matching index*.

The theory behind Hierarchical fuzzy pattern matching is fully described by Power *et al.* (2001). Examples are given by Hagen (2002a). The method is also incorporated in the MAP COMPARISON KIT. However, we have not analysed the Fuzzy pattern matching approach in this report because the method has not yet fully evolved. Moreover, it is not clear what the specific advantages and disadvantages are between the fuzzy-set mapping approach introduced by Hagen (2003) and the method of Power *et al.*.

## 2.6 Comparing maps by use of single map statistics

All methods described in sections 2.2, 2.3 and 2.4 involved map comparison on the basis of *grid-cell similarities*. However, there are many map statistics more conceivable than those based on grid cells. In fact, any statistic that can be calculated for one specific map, can be used in map comparison as well!

For example, if we have a map index for the fragmentation of nature areas, we can compare this fragmentation index across a number of simulated land-use maps for the year 2020. Another application is for the process of calibration. We may want a land-use simulation for the year 1996 to be as similar as possible to a 1996 real land-use map with respect to fragmentation.

There are many map statistics available. We refer to Cressie (1993), Deutsch and Journel (1998), Goovaers (1997) and Kaluzny et al. (1998). However, the number of techniques decreases if we move from maps on the interval/ratio scale to ordinal maps to nominal maps.

We mention four statistics below that grasp important map characteristics; these statistics are specially suited for nominal maps:

- Estimation of spatial trends for two-category maps (any multi-category map can be transformed to a two-category map). Methods available are LOESS estimators, kernel estimators or Indicator Kriging. As in time-series analysis it may be of interest to compare spatial trends between a set of maps, or their residuals, i.e. the data minus the spatial trend. In this way both low and high frequency patterns may be compared across maps. Estimation is only possible for two-category-maps. For example, if one map has a pronounced north-south gradient in the category 'industrial', does another map have this gradient too, or if so, does it have the same slope?
- Estimation of fractal forms of categories within a map. The fractal forms are a measure of the *fragmentation* of a specific category (White and Engelen, 1993). Fragmentation is an important environmental issue for nature and forest areas. For the maintenance of many animal species fragmentation of lands is bound to certain limits.
- Estimation of spatial autocorrelation (the spatial equivalent of the autocorrelation function in time-series analysis). This can be done by standard geostatistical methods (Cressie, 1993) or spatial statistics such as Moran's I statistics for spatial autocorrelation (Anselin, 1988; Verburg *et al.* 2003).

- Estimation of neighbourhood characteristics. These characteristics allow us to quantify the influence of a specific land-use type on other land-use types in the neighbourhood of that location. Such an analysis will allow us to identify how land-use types interact and to judge over which distances such interactions are relevant. Neighbourhood characteristics are important for the calibration of land-use models such as the Environment Explorer, because these characteristics are very similar to distance-decay functions that are an integral part of cellular automata models.

In Appendix A we describe the spatial structure of land use by characterizing the influence of the neighbourhood of a location on its land use. This description is taken from Verburg *et al.* (2003).

### 3. Map Comparison - software

by H. Visser and A. Hagen

#### 3.1 MAP COMPARISON KIT

##### 3.1.1 The software

The MAP COMPARISON KIT has been designed for analysis and comparison of raster maps. **Figure 10** shows a number of screens. Maps may be on nominal, ordinal or ratio scales. Mixtures of nominal and ordinal scales can be also handled by the software (compare: section 2.1). Besides a number of comparison algorithms, the MCK also offers advanced options for visualizing, organizing and exporting raster maps.

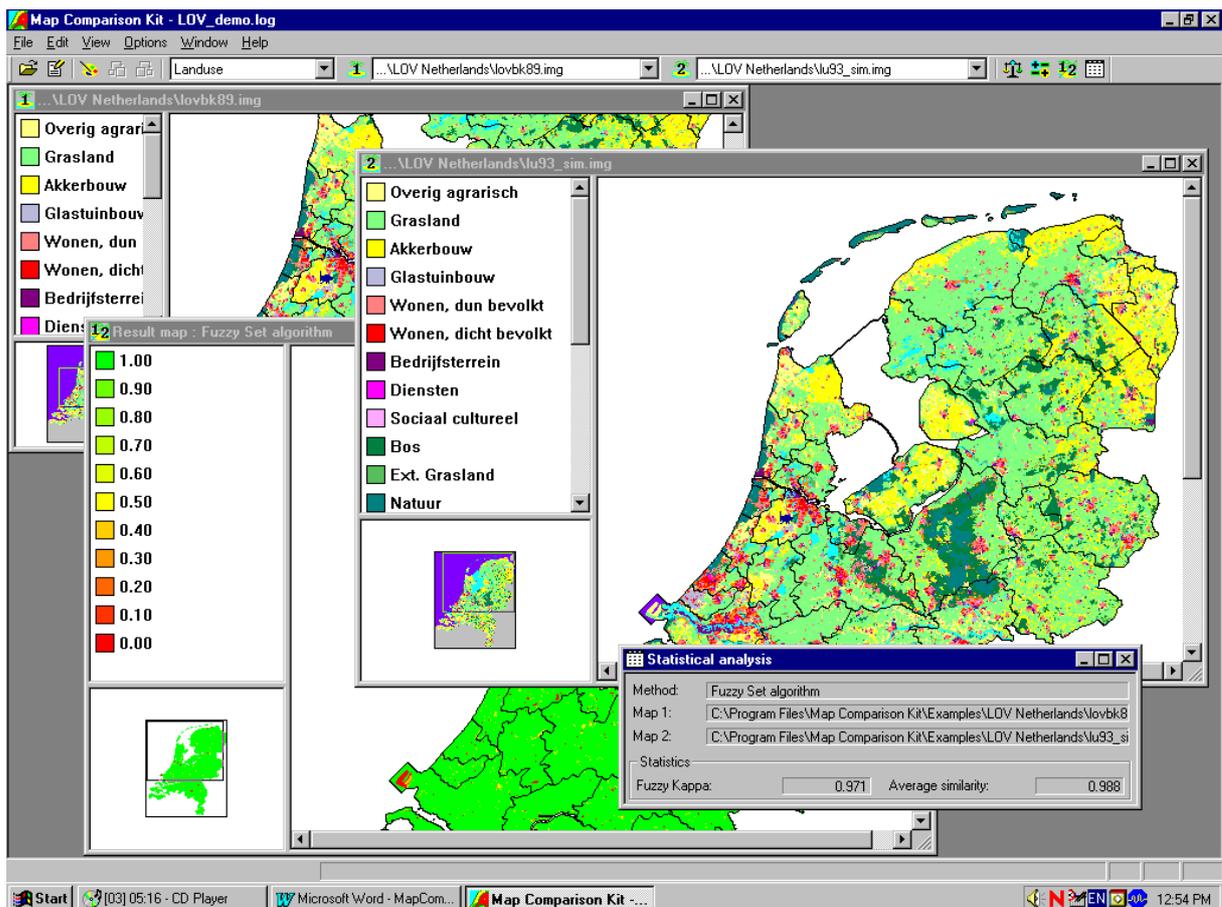


Figure 10 Impression of the MCK screen. Shown are two maps from the RIVM Environment Explorer, a fuzzy result map and corresponding statistical results.

The first version of the MAP COMPARISON KIT dates back to 1992, when it was called the *Analyse Tool*. The software was initially intended for the analysis of series of maps generated as output by simulation software from the Research Institute for Knowledge Systems (RIKS). From 1992 onwards, the tool was steadily further developed as part of the RIKS projects for RIVM, RIKZ, RWS, EC-JRC and others.

The latest extension of the MAP COMPARISON KIT has been developed by order and for the account of RIVM, within the framework of project S/50002/01/TO, Measuring and Modelling. New additions are (i) the extended Kappa analysis, (ii) Hierarchical fuzzy pattern matching and (iii) the Fuzzy-set map comparison. All these map comparison methods are the result of research performed by RIKS, in collaboration with RIVM.

Another novelty is that the software contributes added value by not only being suited to work in conjunction with other RIKS products, but may also be used to compare any set of raster maps in some of the most popular file formats. Specifically, these are ArcInfo ASCII format, Idrisi Raster format and the LLO format used at RIVM. A third novelty is the simple import of maps and statistical results to WORD files, as shown in Figure 10. The active screen is copied to the Windows clipboard using 'Ctrl C' and pasted into the Word file with 'Ctrl V'.

These extensions have stretched the use of the tool beyond the analysis of RIKS simulation results, and inspired the name change from *Analyse Tool* to *MAP COMPARISON KIT*. The software is available on CD as a stand-alone package. The CD also contains a user manual of which Alex Hagen is the lead author (RIKS, 2003a). As part of the software a number of map datasets have been added to ease understanding of the software. The software will also become available on Internet at the beginning of 2004.

Besides the MCK stand-alone version, a Raster Comparison Module containing three comparison methods has been programmed: (i) cell-by-cell, (ii) Hierarchical fuzzy pattern matching and (iii) the Fuzzy-set algorithm. The module can be incorporated into GIS packages such as ArcGis-8. The reference manual for the Map Comparison Module is from I. Uljee (RIKS, 2003b). A simple prototype user shell in ArcGis-8 has been programmed by G.J. Verspaaij and G.J.C. Stolwijk (both IMP/MNP).

### 3.1.2 Handling different measurement scales

**Nominal maps** may be compared cell by cell or by fuzzy methods. The MCK screen for choosing one of these methods is given in **Table 4**.

In the following sections we will exploit these methods for analysing: (i) children's maps (section 4.1), (ii) historical changes in agricultural land use (section 4.2) and (iii) the analysis of different land-use scenarios for the year 2020 (section 4.3). Nominal maps are also analysed in Chapter 5.

Table 4 MCK screen for choosing a nominal comparison method.

methods for nominal  
map comparison

The screenshot shows a dialog box titled "Comparisons and other operations" with a close button (X) in the top right corner. The dialog is divided into three main sections:

- Compare categories:** This section contains four radio button options:
  - Per category
  - Cell by cell
  - Fuzzy Inference System algorithm
  - Fuzzy Set algorithm
- Compare numerical values:** This section contains seven radio button options:
  - second - first
  - abs( second - first )
  - ( second - first ) / max( abs( second - first ))
  - abs( second - first ) / max( abs( second - first ))
  - second / first
  - abs( second / first )
- NODATA:** This section contains two radio button options:
  - f(NODATA,x) = NODATA
  - f(NODATA,x) = f(0,x)
- Other operations:** This section contains three radio button options, each with a text input field:
  - first + second
  - first +
  - first \*

On the right side of the dialog, there are two buttons: "OK" and "Cancel".

The **ordinal** character of a set of maps can be exploited using the so-called Category Similarity Matrix. In standard use, this matrix is unitarian, meaning that each category is defined similarly in both maps (ones on the main diagonal), and different categories have no overlap (zeros at all off-diagonal places) (see **Table 5**). In ordinal maps this situation typically not the case. If we have the two categories of ‘sparse residential’ and ‘dense residential’, we may want to make these similar by applying a value of, say, 0.5. A value of 1.0 would make both categories totally equal and would in fact define the new category ‘total residential’. Since the choice of a particular value is subjective, it has to be chosen on the basis of a-priori insights.

In section 4.4 we will use maps from the Land Use Scanner to illustrate the procedure.

*Table 5      Advanced settings for fuzzy calculations within the MCK software.*

**Fuzzy Kappa Settings**

Settings Map 1

Radius of neighbourhood: 4

Exponential decay  
Halving distance: 2

Linear decay  
Slope: 0.5

Constant value  
Value: 0.5

Settings Map 2

Radius of neighbourhood: 4

Exponential decay  
Halving distance: 2

Linear decay  
Slope: 0.5

Constant value  
Value: 0.5

Buttons: OK, Cancel, Apply, Standard <<

Category Similarity Matrix

Map 1 \ Map 2	no r...	0 to ...	10 t...	20 t...	30 t...	40 t...	50 t...	60 t...	70 t
no residential	1	0	0	0	0	0	0	0	0
0 to 10% residential	0	1	0	0	0	0	0	0	0
10 to 20% residential	0	0	1	0	0	0	0	0	0
20 to 30% residential	0	0	0	1	0	0	0	0	0
30 to 40% residential	0	0	0	0	1	0	0	0	0
40 to 50% residential	0	0	0	0	0	1	0	0	0
50 to 60% residential	0	0	0	0	0	0	1	0	0
60 to 70% residential	0	0	0	0	0	0	0	1	0
70 to 80% residential	0	0	0	0	0	0	0	0	1
80 to 90% residential	0	0	0	0	0	0	0	0	0
90 to 100% residential	0	0	0	0	0	0	0	0	0

Buttons: Load..., Save..., Unity

When we are dealing with maps **on interval or ratio scales**, we can choose two paths of analysis. First, we can transform the continuous variable to disjunct categories and treat the maps subsequently as ordinal. This procedure is followed in section 4.4 for ascertaining ‘percentage residential’. With use of the MapTransformer software we generated 11 ordinal categories.

Second, we can exploit the character of specific continuous variables and subtract or divide values for identical grid cells. We then get result maps that show map differences on a *continuous* scale. The MCK screen for choosing the calculation of such result maps is shown in **Table 6**. Examples are given in sections 4.5 and 4.6.

Table 6 MCK screen for manipulations on a set of ‘continuous’ maps.

methods for comparing maps on interval or ratio scales

The screenshot shows a dialog box titled "Comparisons and other operations" with the following options:

- Compare categories:**
  - Per category
  - Cell by cell
  - Fuzzy Inference System algorithm
  - Fuzzy Set algorithm
- Compare numerical values:**
  - second - first
  - abs( second - first )
  - ( second - first ) / max( abs( second - first ))
  - abs( second - first ) / max( abs( second - first ))
  - second / first
  - abs( second / first )
- NODATA:**
  - f(NODATA,x) = NODATA
  - f(NODATA,x) = f(0,x)
- Other operations:**
  - first + second
  - first +
  - first \*

Buttons: OK, Cancel

## 3.2 Tools additional to the MAP COMPARISON KIT

There are a number of software tools for reading, generating and transforming maps, which can be used in conjunction to the MCK software. Especially, the import and export facilities of the MCK with respect to the widely used ArcInfo ASCII files ease such combinations. We name here ArcView, Idrisi and Fragstats. RIVM environmental modelling software, such as the Environment Explorer, Nature Explorer and Land Use Scanner, is also used easily in combination with the MCK.

Furthermore, one can easily program input and output routines in well-known software environments such as Fortran, S-PLUS or Mathematica.

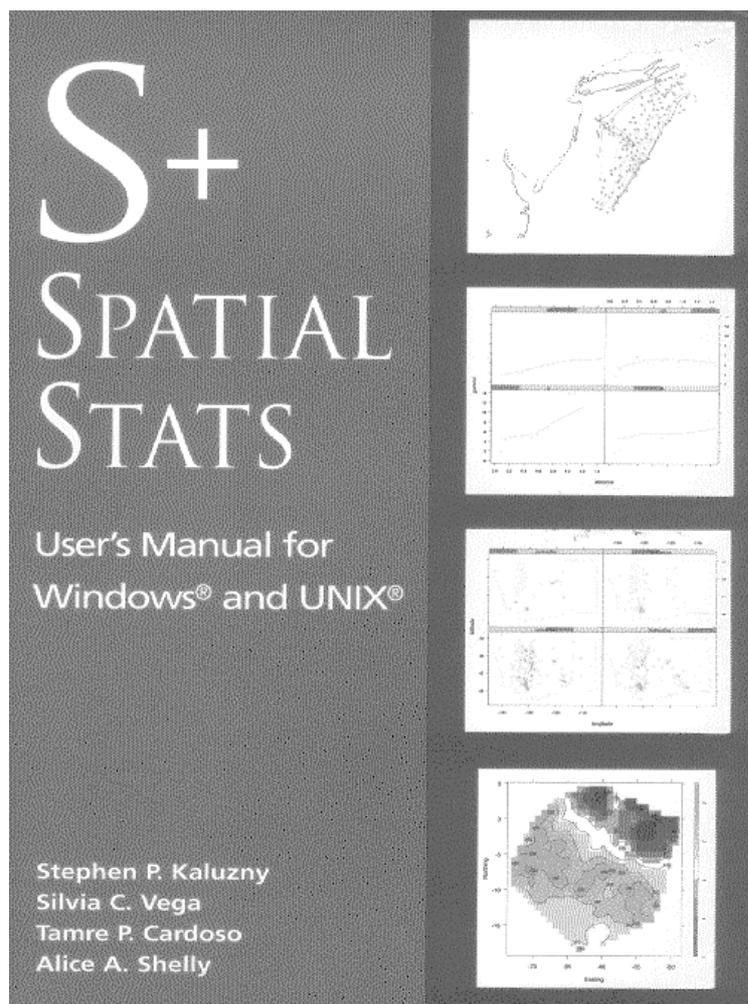
We will later describe two additional tools in **Appendix B**, both developed at RIVM. The first tool performs transformations to a set of maps in ArcInfo ASCII files, while the second tool enables the import and export of maps in ArcInfo ASCII format to S-PLUS.

## 3.3 Software for the calculation of additional statistics

Many of the map statistics named in section 2.6 can be calculated through the use of the SpatialStats module from S-PLUS. We refer to the textbook by Kaluzny *et al.* (1998).

Software, called INFLUENCE, for calculating enrichment factors (Appendix A) has been developed by Kor de Jong of Utrecht University. We note that INFLUENCE is able to cope with maps that have missing data. The source code and executables can be downloaded from the project homepage: <http://networks.geog.uu.nl>.

Maps in the ArcInfo ASCII grid format (unfortunately) cannot be imported directly to INFLUENCE. To import these maps, one has to import maps in the software package PCRaster, after which the maps are exported in the PCRaster CSF-2.0 format. The PCRaster package can be downloaded from <http://www.pcraster.nl> and <http://pcraster.geog.uu.nl>.



*A textbook written by Kaluzny et al. (1998) contains a great number of statistics for spatial data and a number of input – output routines for GIS applications. The SpatialStats module forms an integral part of the S-PLUS software at RIVM (Dekkers, 2001).*



## 4. MAP COMPARISON KIT – applications

Here, we will give six examples of map comparison applications. The first example is a humouristic illustration for two children's maps, both on a cell-by-cell and a fuzzy-set theory basis. Sections 4.2, 4.3, 4.4 and 4.6 give case studies from the daily NEAA practice. Section 4.5 on ambulance driving times is taken from the Environmental Health group (VTV) of RIVM.

The examples comprise all types of maps as described in section 2.1: nominal maps (sections 4.1, 4.2 and 4.3), ordinal maps (section 4.4), and interval/ratio maps (sections 4.5 and 4.6).



*The prediction of land use 20 to 30 years in advance is an important task of NEAA. This is done on the basis of scenarios developed for future economic and demographic growth. These scenarios function as input for land-use models such as the Environment Explorer and the Land Use Scanner. Of special importance is the 'tension' between growing urbanization on the one hand, and environmental care for nature/forest areas on the other hand.*

*Photo: Advertisement campaign – Gemjobs.*

## 4.1 Spot the 15 differences!

by Hans Visser and Alex Hagen

### 4.1.1 Introduction

We will illustrate the practical use of Fuzzy-set map comparison versus cell-by-cell through the use of two hobbit drawings (maps) given in **Figure 11**. The task for the reader is to spot 15 differences in the ‘maps’. Before reading further, it is helpful to find these differences and mark them with a pencil.

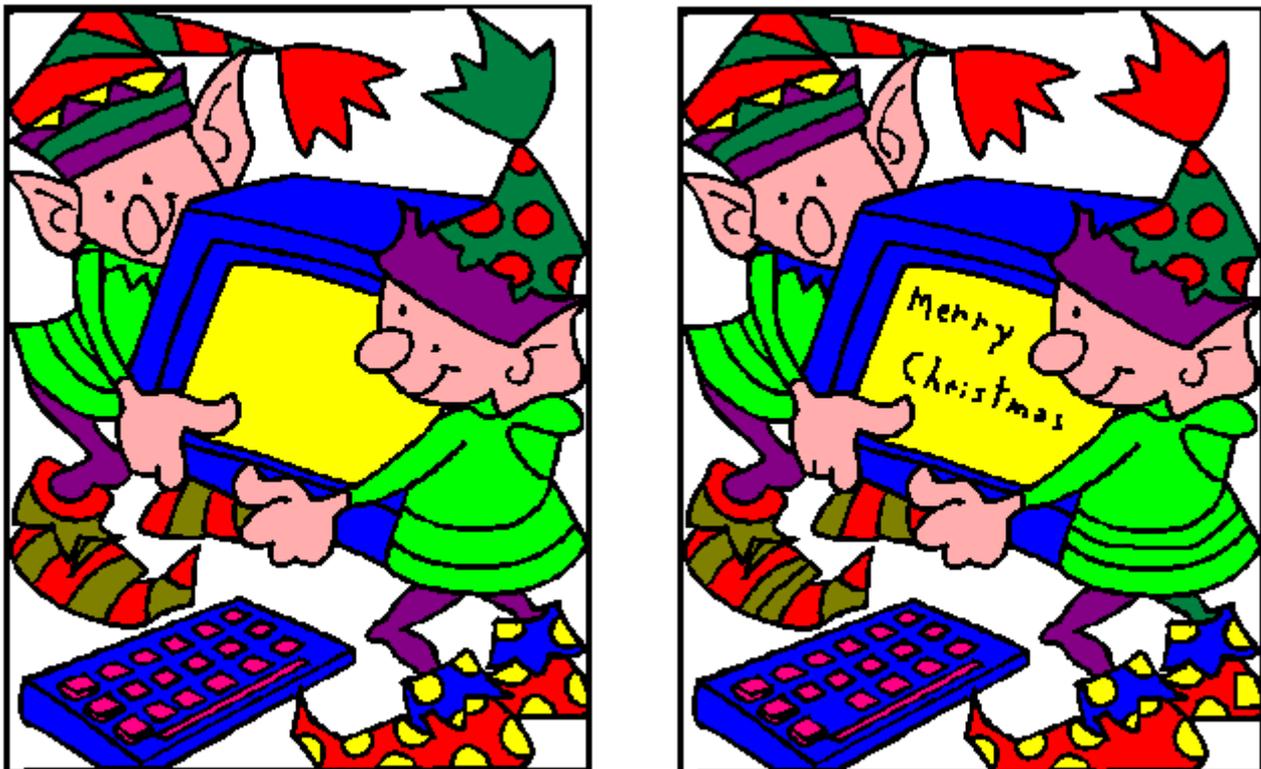


Figure 11 Two hobbit drawings. The task is to spot 15 differences in the drawings (‘maps’).

### 4.1.2 Cell-by-cell comparison

The cell-by-cell, equal-unequal, map is given in **Figure 12**. The result map shows approximately 10 differences and a great number of red contour lines, originating from the fact that we have shifted the original second hobbit map a few cells to the left. Such a shift is not seen by the human observer. The shift led to the red contour lines around the hobbits and the PC. Such shifts or distortions often occur in practical situations. They can be due to re-projection of data, gridding and re-gridding, or vague transitions between categories.

The Kappa statistics are summarized in **Table 7**. The Kappa for the colour black seems to be lowest (0.58), an obvious result of the complete shift of the second drawing.

However, we are not interested in these shifts. We only want to find the 15 differences. Five differences are masked in Figure 12 by the contour lines. There has to be a better method.

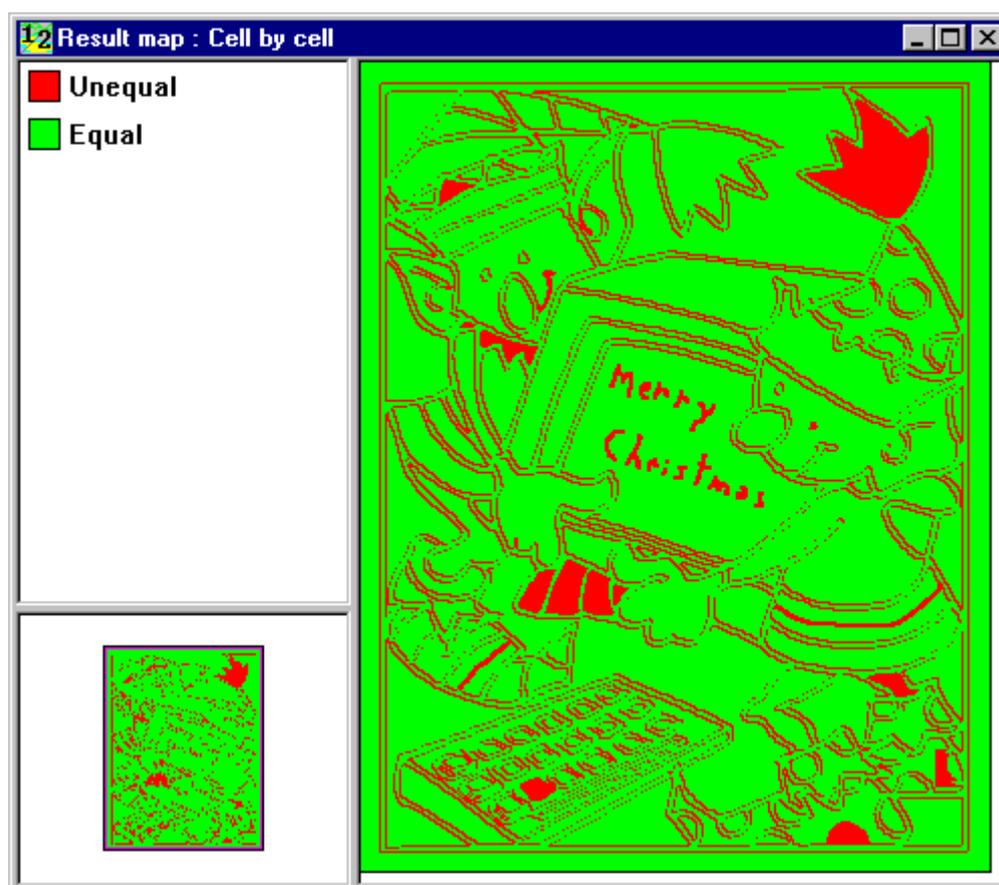


Figure 12 Equal-unequal result map.

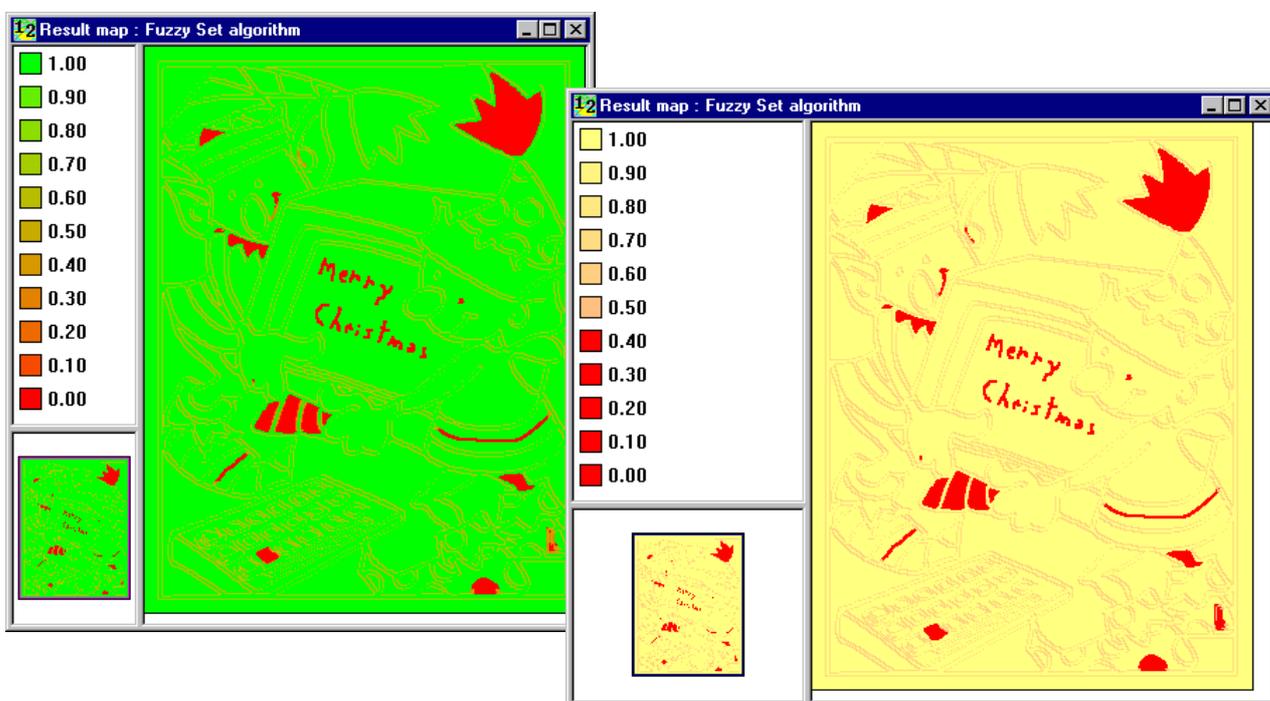
Table 7 Kappa statistics for the hobbit drawings.

Statistics													
Kappa:	0.801			Fraction correct:	0.836								
KLocation:	0.824			KHisto:	0.972								
Per Category													
	black	brown	dark green	green	yellow	skin color	white	pink	purple	red	blue		
Kappa	0.581	0.652	0.643	0.890	0.901	0.906	0.907	0.690	0.886	0.713	0.846		
KLoc	0.591	0.658	0.813	0.911	0.936	0.908	0.907	0.712	0.909	0.811	0.859		
KHisto	0.983	0.991	0.791	0.976	0.963	0.998	1.000	0.970	0.974	0.879	0.985		
Contingency Table													
Map 1 \ Map 2	black	brown	dark green	green	yellow	skin color	white	pink	purple	red	blue	Nodata	Sum Map1
black	18165	209	373	718	497	917	2536	262	388	981	1662	0	26708
brown	291	1162	0	0	0	0	0	0	0	331	0	0	1784
dark green	440	0	2494	0	0	0	0	0	0	1659	0	0	4593
green	981	0	0	7976	0	0	0	0	0	0	126	0	9083
yellow	1063	0	0	0	8345	0	0	0	0	104	0	0	9512
skin color	888	0	0	0	0	9594	0	0	0	0	0	0	10482
white	2536	0	0	0	0	0	36917	0	0	0	0	0	39453
pink	261	0	0	0	0	0	0	655	0	0	58	0	974
purple	438	0	176	0	0	0	0	0	4056	0	0	0	4670
red	839	382	0	0	31	0	0	0	0	5866	0	0	7118
blue	1526	0	0	0	0	0	0	0	0	0	10457	0	11983
Nodata	0	0	0	0	0	0	0	0	0	0	0	0	0
Sum Map 2	27428	1753	3043	8694	8873	10511	39453	917	4444	8941	12303	0	126360

### 4.1.3 Fuzzy-set map comparison

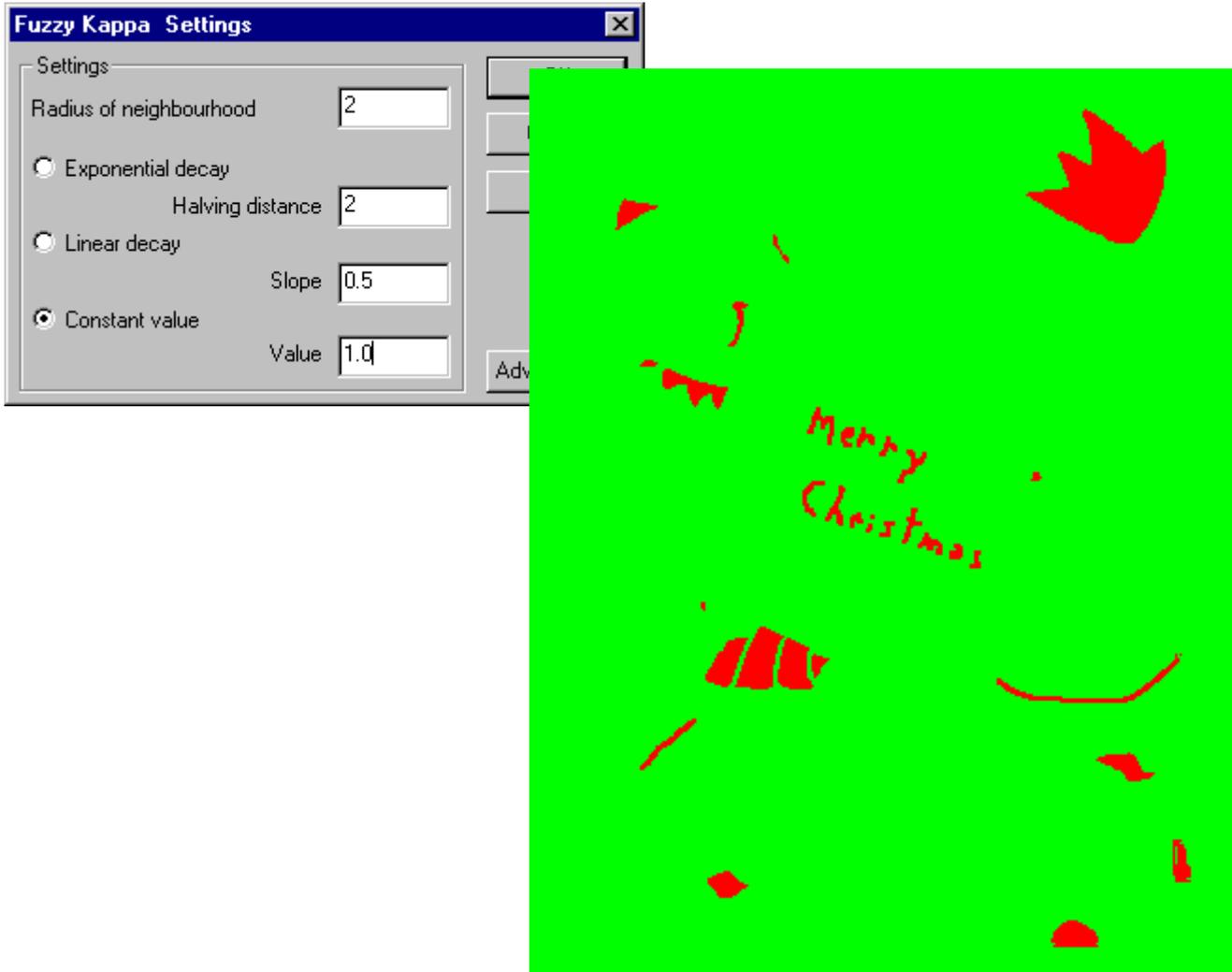
**Figure 13 (left panel)** shows the result map following the fuzzy-set approach (see section 2.4), using the default settings of the MCK software. Because the black contour lines shifted only a few cells, the fuzzy-result map shows these in the yellow colour. These lines are now less pronounced.

It is easy to find 14 differences in the left panel in Figure 13. However, we want to find the 15th difference too. To this end, we changed the colouring in the legend. The result is shown in the **right panel of Figure 13**, where all differences are easily identified.



*Figure 13 Fuzzy-set result map for the hobbit maps shown in Figure 11. Left and right panels differ only in the colouring of the legend.*

If we want to remove the contour lines from the presentations in Figure 13, we can choose another distance function within the fuzzy context: a constant distance with a weighting factor 1.0. For the maximum distance ('radius of neighbourhood') we chose two cells (the left panel in Figure 11 shifted by only two cells). The result is shown in **Figure 14**, where exactly the 15 differences and no contour lines are visible.



*Figure 14 By choosing the distance function 'Constant value' with a value of 1.0, and a neighbourhood radius of 2, we spot the 15 differences exactly without any contour lines. Merry Christmas!*

#### **4.1.4 Conclusion**

What can be learnt from these children's 'maps'? Firstly, for serious maps we are often interested in distinguishing between small differences in location and real differences in location. Clearly, it is the Fuzzy-set map comparison that makes the difference (Figures 13 and 14).

Secondly, shifts or small distortions in the forms of cells often occur in practice. These differences originate from the application of different cartographic projections and/or different sampling techniques (e.g., remote sensing versus ground-based inventories). However, in analysing the (dis-)similarities between two maps, we are interested in real differences and not in these sampling artefacts. Fuzzy techniques are able to make such distinctions.

## 4.2 Historical projections of agricultural land use, 1700-1990

by C.G.M. Klein Goldewijk and H. Visser

### 4.2.1 Context

Since the beginning of civilization, humans have managed and adapted the landscape to derive valuable natural resources such as food, fiber, fresh water and pharmaceutical products. As early as 1864, Marsh voiced concern about the deleterious consequences of human activities. Richards has estimated that over the last three centuries, roughly 1200 million ha of forests and woodlands have been cleared; grasslands and pastures have diminished by about 560 million ha (however, many grasslands have been converted to pastures) and cropland areas have increased by 1200 million ha. At present, land totalling about 1800 million ha, an area roughly the size of South America, is used for some form of cultivation (Klein Goldewijk and Ramankutty, 2003).

Such large-scale changes in land use and land cover can have significant biophysical and biogeochemical consequences. For example, changes in surface reflectivity and water balance associated with land cover change can modify regional and global climates. Land use change over the past 150 years has released almost half as much carbon as that from fossil fuel emissions. Postel and Vitousek have reviewed the consequences of human activities for the global cycles of water and nitrogen. The emerging understanding of the consequences of land-use and land-cover change has led to a growing demand for global spatially explicit historical data on land use/land cover. Climate and ecosystem models, which need such spatially and temporally explicit data as inputs, can then be used to increase our understanding of the consequences for the Earth system.

### 4.2.2 Data

Two research groups have developed land-use and land cover data sets on global scale. Firstly, Ramankutty and Foley (1998) used the IGBP 1 km-resolution Global Land Cover Classification (GLCC) data set and calibrated it against cropland inventory data for 1992 to create a continuous global data set of cultivated land for 1992. They used a 'hindcast' modelling technique to extrapolate this data backwards by combining it with a compilation of historical cropland inventory data to create a data set of croplands from 1700 to 1992. Ramankutty and Foley also created a global potential natural vegetation dataset of 15 different vegetation types by combining the GLCC data set with the Haxeltine and Prentice potential vegetation map. By overlaying the croplands dataset with the potential vegetation dataset, one can estimate the extent to which different vegetation types have been converted to cultivation.

Henceforth, this dataset is referred to as the ‘RF’ data set (available from <http://www.sage.wisc.edu>).

Secondly, Klein Goldewijk’s (2001) work involved the construction of a History Database of the Global Environment, henceforth referred to as the ‘HYDE’ data set. HYDE was originally designed for testing and validating the IMAGE model, because international reviews of IMAGE suggested that if the model is to make projections for the next hundred years, it should at least be able to simulate the last century).

Klein Goldewijk used the potential vegetation dataset of RF as a starting point, and, besides the natural vegetation, also allocated the historical inventory data on agricultural activities like cropland and pasture with historical population density maps as a proxy for location. The assumption was made that where people had lived for many centuries, agricultural activity was more likely to be in the proximity of urban centres than in remote areas. A boolean approach was chosen: each grid cell was, for simplicity, totally allocated to cropland, pasture or one of the potential land-cover classes.



*Projections of historical agricultural land use were generated by starting at present land use (1992) and by using hindcast modelling techniques to extrapolate these data back to the year 1700.  
Photo: H. Visser.*

The historical population density maps were created by downscaling the population density map of Tobler *et al.* (1995) on the basis of historical total population numbers for (sub-) administrative units (the largest unit being country level). Grid cells with the highest historical population density were classified as being most likely available for agricultural expansion at that time. Thus, total population numbers per country (or per sub-administrative unit) are consistent with the statistics, while the density patterns reflect the current pattern. Reasoning for this approach is that towns/cities did not move spatially during the last 300 years. Thus, the distribution of the current density map is also valid for the historical time period. The approach to this study is described in detail in Klein Goldewijk (2001). Results are available at the website <http://arch.rivm.nl/env/int/hyde>.

**Figure 15A** shows the three maps of Ramankutty and Foley (RF) for 1700, 1850 and 1990. **Figure 15B** shows comparable maps of Klein Goldewijk (HYDE).

### 4.2.3 Similarity based on Kappa statistics

How similar are the RF and HYDE maps? And how is this (dis)similarity influenced by the projection time interval? A first impression is given in **Table 8**, showing Kappa statistics. As expected, Kappa statistics decrease as projections go back further in time. For HYDE the  $Kappa_{1850,1990}$  is 0.51 and the  $Kappa_{1700,1990}$  is 0.32 (green figures in the table).

The Kappas between the corresponding RF and HYDE maps also decrease with time:  $Kappa_{1990} = 0.47$ ,  $Kappa_{1850} = 0.33$  and  $Kappa_{1700} = 0.20$  (red figures in the table). All Kappas are fairly low. Especially for 1990, one would have expected a higher Kappa value.

*Table 8 Kappa's for HYDE and RF maps.*

<b>Kappa</b>	<b>Hyde 1700</b>	<b>Hyde 1850</b>	<b>Hyde 1990</b>	<b>RF 1700</b>	<b>RF 1850</b>	<b>RF 1990</b>
<b>Hyde 1700</b>	1.00	0.68	0.32	<b>0.20</b>	-	-
<b>Hyde 1850</b>	0.68	1.00	0.51	-	<b>0.33</b>	-
<b>Hyde 1990</b>	<b>0.32</b>	<b>0.51</b>	<b>1.00</b>	-	-	<b>0.47</b>
<b>RF 1700</b>	0.20	-	-	1.00	0.62	0.32
<b>RF 1850</b>	-	0.33	-	0.62	1.00	0.56
<b>RF 1990</b>	-	-	0.47	0.32	0.56	1.00

Where do these map differences originate? **Table 9** shows the K-location and K-histo values for HYDE and RF for all relevant map combinations. High K-location values point to the fact that most cells for agriculture in both maps are at the same locations. High K-histo values

point to the fact that the total number of cells for agriculture are in very close proximity in both maps. The K-location values in the table show that locations of agriculture did not move for either HYDE or RF maps. For HYDE we find  $K\text{-location}_{1990,1850} = 0.96$  and  $K\text{-location}_{1990,1700} = 1.00$  (green figures). Here, we see a typical characteristic of both projection methods: cells for agriculture in 1700 did not change positions over the 300 year period! Only the frequency changed over the years (K-histo values low).

Table 9 K-location (upper panel) and K-histo (lower panel) for HYDE and RF maps.

K-location	Hyde 1700	Hyde 1850	Hyde 1990	RF 1700	RF 1850	RF 1990
Hyde 1700	1.00	0.99	1.00	<b>0.26</b>	-	-
Hyde 1850	0.99	1.00	0.96	-	<b>0.43</b>	-
Hyde 1990	<b>1.00</b>	<b>0.96</b>	<b>1.00</b>	-	-	<b>0.55</b>
RF 1700	0.26	-	-	1.00	1.00	0.98
RF 1850	-	0.43	-	1.00	1.00	0.93
RF 1990	-	-	0.55	0.98	0.93	1.00

K-histo	Hyde 1700	Hyde 1850	Hyde 1990	RF 1700	RF 1850	RF 1990
Hyde 1700	1.00	0.69	0.32	<b>0.81</b>	-	-
Hyde 1850	0.69	1.00	0.52	-	<b>0.76</b>	-
Hyde 1990	<b>0.32</b>	<b>0.52</b>	<b>1.00</b>	-	-	<b>0.87</b>
RF 1700	0.81	-	-	1.00	0.62	0.33
RF 1850	-	0.76	-	0.62	1.00	0.61
RF 1990	-	-	0.87	0.33	0.61	1.00

Comparison between HYDE and RF shows  $K\text{-location}_{1990} = 0.55$ ,  $K\text{-location}_{1850} = 0.43$  and  $K\text{-location}_{1700} = 0.26$  (red figures in upper panel of the table). These figures are very similar to the corresponding Kappas, meaning that differences between HYDE and RF are mainly due to differences in location. This is confirmed by the high K-histo values: 0.81, 0.76 and 0.87 (red figures in the lower panel of the table).

Hence, we conclude that HYDE and RF maps differ according to agricultural locations. However, because all Kappas are calculated on a cell by cell basis, we cannot say anything about the exact nature of these location differences. This point is studied in more detail in the next section.

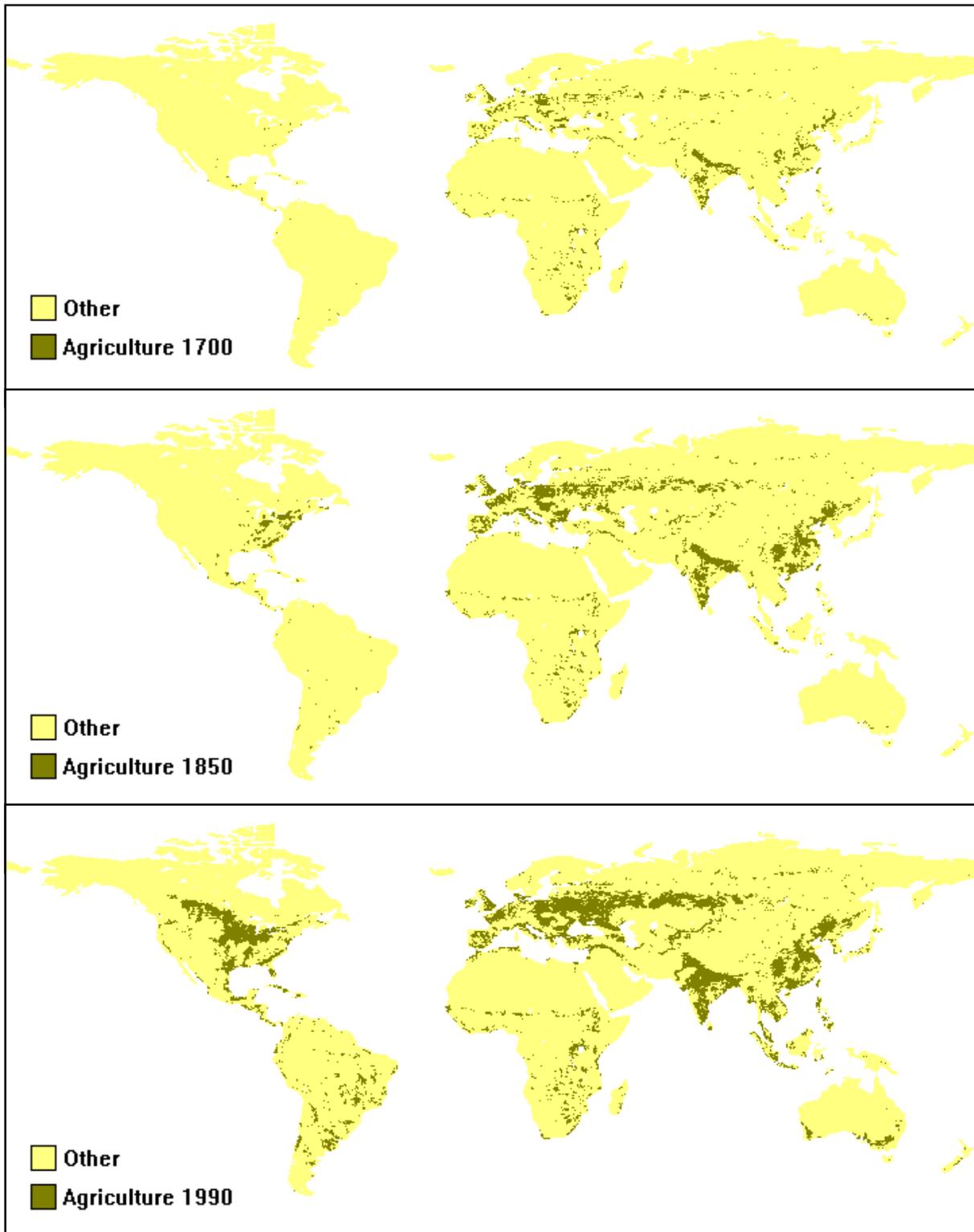


Figure 15A RF maps for the years 1700 (upper panel), 1850 (middle panel) and 1990 (lower panel).

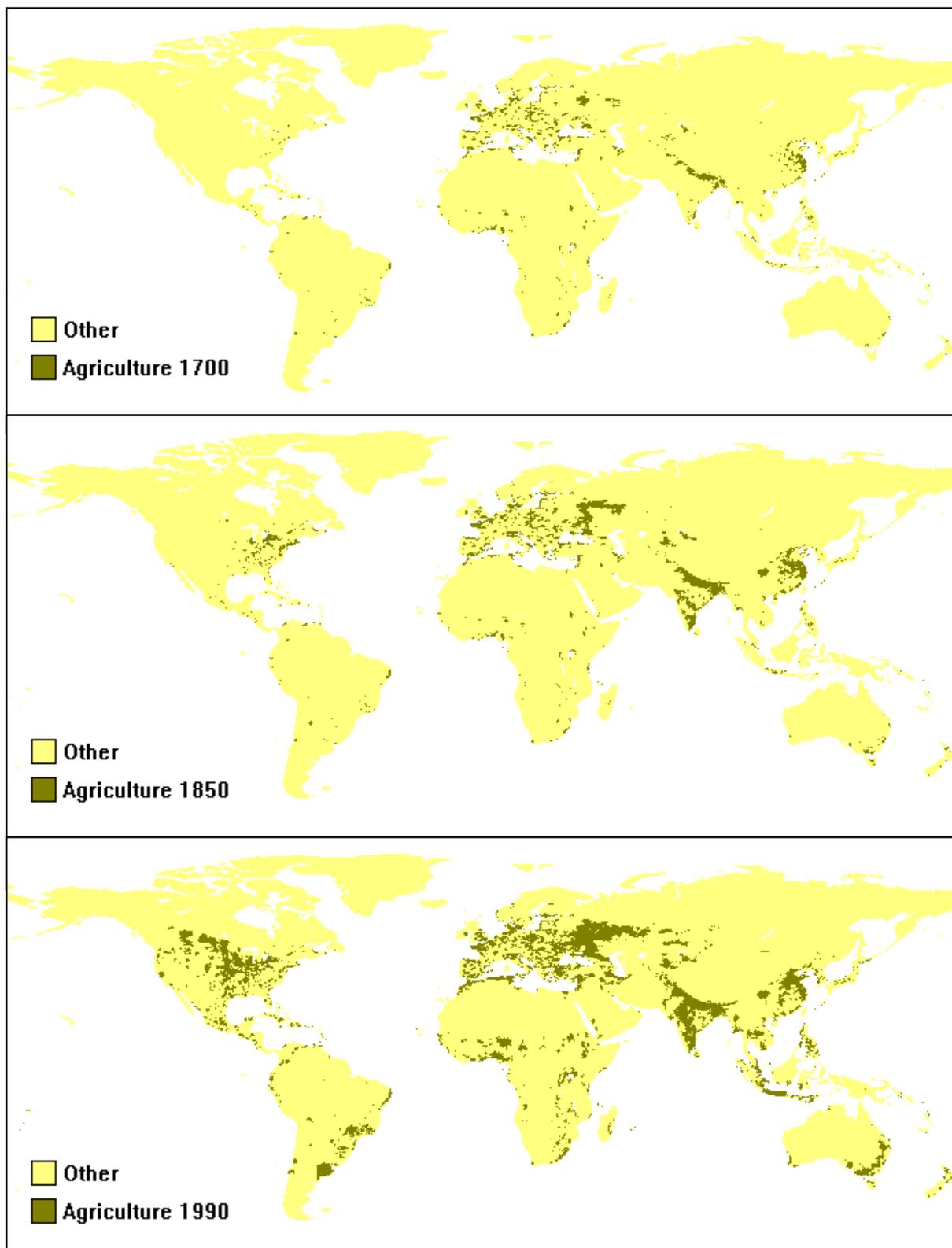


Figure 15B HYDE maps for the years 1700 (upper panel), 1850 (middle panel) and 1990 (lower panel).

#### 4.2.4 Fuzzy-set map comparison

We applied the fuzzy-set approach to study location differences (section 2.4). The overall K-fuzzy values are given in **Table 10**. The differences between Kappa (given in Table 8) and Fuzzy Kappa appear to be small. The only difference is that the Fuzzy Kappa between RF and HYDE remains constant for all years (red figures).

*Table 10 Fuzzy Kappa estimates using the default values of the Fuzzy- set map comparison approach.*

<b>Fuzzy Kappa</b>	<b>Hyde 1700</b>	<b>Hyde 1850</b>	<b>Hyde 1990</b>	<b>RF 1700</b>	<b>RF 1850</b>	<b>RF 1990</b>
<b>Hyde 1700</b>	1.00	0.73	0.25	<b>0.28</b>	-	-
<b>Hyde 1850</b>	0.73	1.00	0.41	-	<b>0.28</b>	-
<b>Hyde 1990</b>	<b>0.25</b>	<b>0.41</b>	<b>1.00</b>	-	-	<b>0.31</b>
<b>RF 1700</b>	0.28	-	-	1.00	0.70	0.25
<b>RF 1850</b>	-	0.28	-	0.70	1.00	0.44
<b>RF 1990</b>	-	-	0.31	0.25	0.44	1.00

**Figure 16** shows a detailed comparison between HYDE and RF. The figure shows the fuzzy-result maps for the years 1700, 1850 and 1990. Green cells point to identical categories in both maps, red cells to totally different categories in both maps and yellow to cells for which categories are nearby in both maps. The great advantage of the Fuzzy-set map comparisons is that we can discriminate between small differences in location (yellow colours) and between real differences (red colours) even on a semi-continuous scale.

For Northern America we see hardly any red cells for 1850 and 1990. Apparently, HYDE and RF are very similar for this continent. For Europe, differences are mainly shown in yellow. Some red occurs for Ireland and Scotland. However, large discrepancies occur in the former Soviet Union.

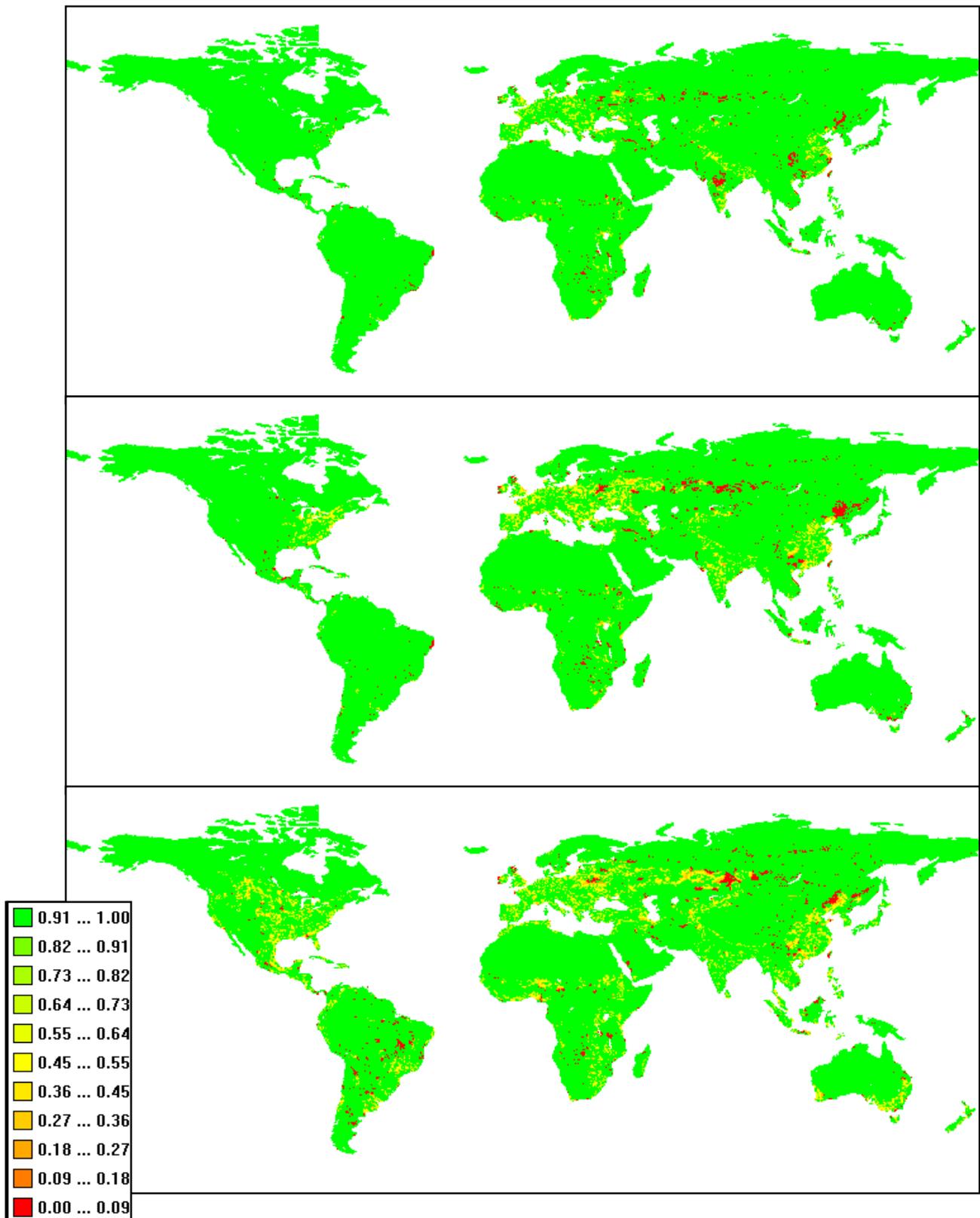


Figure 16 Fuzzy-set map comparisons for 1700 (upper panel), 1850 (middle panel) and 1990 (lower panel). Default values were used for the Fuzzy-set method.

### 4.2.5 Effect of aggregation

Another view on spatial differences is gained if cells are aggregated to larger surfaces (change to larger support). One form of aggregation is easily performed within the fuzzy context: in the Fuzzy set method choose the distant function ‘constant’ with a value of ‘1.0’. As an example we have chosen four levels of aggregation: 0.5 by 0.5 degrees (smallest aggregation level), 2.0 by 2.0 degrees, 4.0 by 4.0 degrees and 6.0 by 6.0 degrees.

Please note that the term ‘aggregation’ should be handled with care here. The form of aggregation here is not proportional to the number of surrounding cells in the other map. There are available methods, based on ‘proportional windows’, which resemble the well-known aggregation to coarser grid maps without losing information. It would be interesting to add such a method to a new version of the MCK software.

The Fuzzy Kappa for these levels of aggregation and different years is given in **Table 11**. The table shows Fuzzy Kappa to strongly increase if the level of aggregation is increased. For instance, for the year 1990 Fuzzy Kappa increases from 0.47 to 0.93. The Fuzzy-set map comparisons for 1990 and three levels of aggregation are shown in **Figure 17**. The upper panel equals the simple cell-by-cell, equal-unequal map. The number of red cells is considerable. However, when moving to a 6.0 by 6.0 degrees aggregation level, almost all red cells have disappeared.

*Table 11 Fuzzy Kappa estimates for three levels of aggregation.*

Aggregation level	Kfuzzy		
	1700	1850	1990
0.5 by 0.5 degrees (distance 0 cells, weight 1.0)	0.22	0.33	0.47
2.0 by 2.0 degrees (distance 4 cells, weight 1.0)	0.55	0.55	0.63
4.0 by 4.0 degrees (distance 8 cells, weight 1.0)	0.69	0.72	0.87
6.0 by 6.0 degrees (distance 12 cells, weight 1.0)	0.82	0.82	0.93

We chose the distance function ‘constant’, with a weight of 1.0. In this way each Fuzzy-set map comparison weights the surrounding cells equally. The first Fuzzy Kappa in the table with a distance of 0 equals the traditional Kappa.

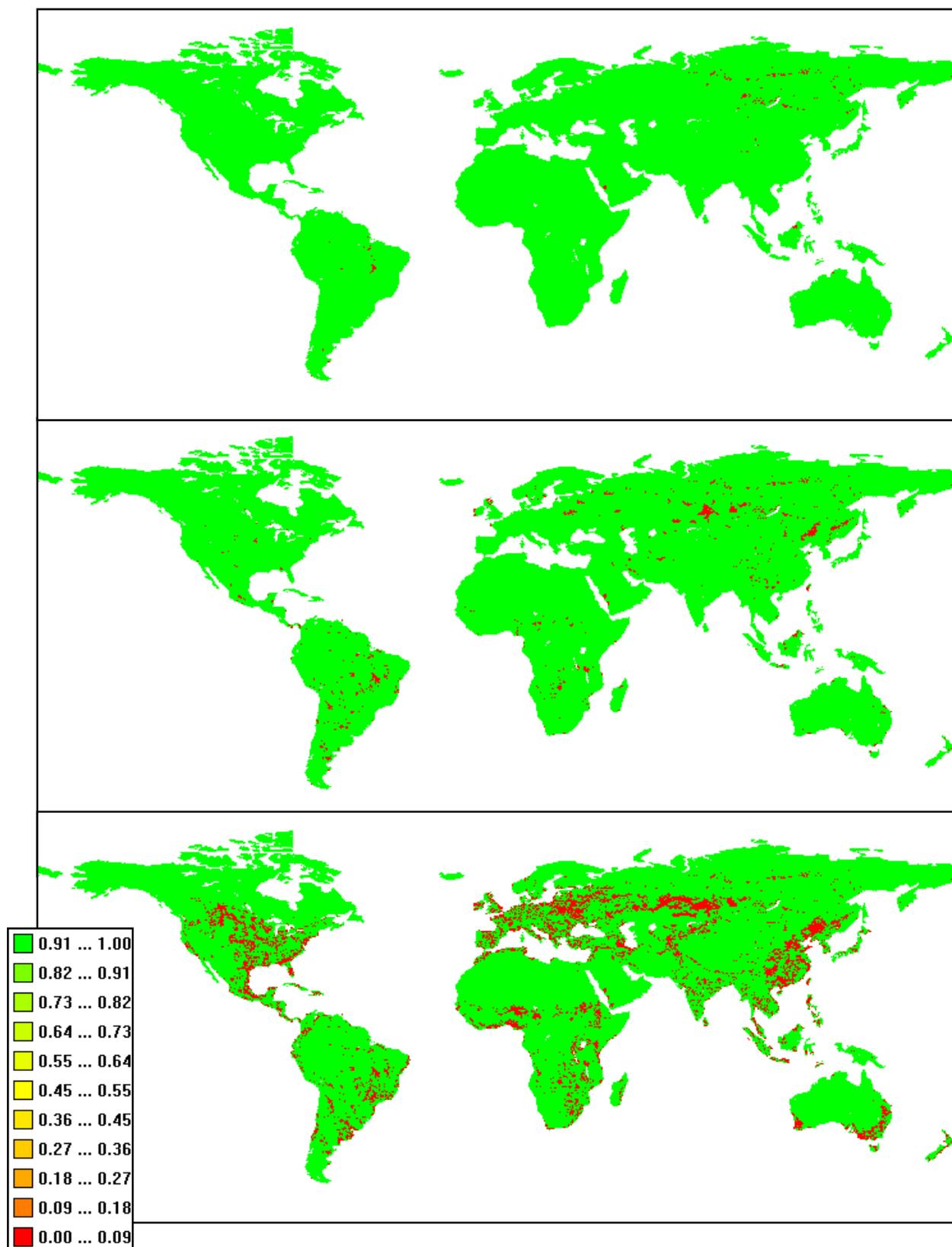


Figure 17 Comparison of HYDE and RF for the year 1990.

Three levels of aggregation are shown: distances of 0 (upper panel), of 4 (middle panel) and of 12 (lower panel). These distances coincide with aggregation levels of 0.5 by 0.5 degrees, 2.0 by 2.0 degrees and 6.0 by 6.0 degrees, respectively.

## 4.2.6 Conclusion

The recognition that global land-use and land cover change is an important driver of global change has been growing. There have been estimates of the magnitude of land cover change at continental scales. Nevertheless, spatially explicit global land-use data sets describing the changes at regular time intervals have, until recently, been absent. Such data sets are critical for evaluating the consequences of land cover change within global climate and ecosystem models. Here, we have examined two new global data sets of historical land-use and land cover change. Both data sets describe the changes in agricultural land from 1700 up to the present.

We have compared the HYDE and RF projections cell by cell (Kappa statistics) and by means of the fuzzy-set theory. The results show a clear advantage of the latter approach. For most continents cell-by-cell comparison yields an underestimation of the neighbourhood. Fuzzy-set map comparison is better, especially with a hindcast result. This is particularly important for climate models, for which these data sets are input. ‘Nearby’ has been considered good enough for these models.

A second advantage is that we are able to discriminate between small differences in location of agriculture and real differences in both historical projections. Such a distinction helps in finding, explaining and correcting inconsistencies between the HYDE and RF models. These inconsistencies can be largely attributed to the following causes:

- 1) use of different assumptions on spatial patterns – RF used remotely-sensed data to spatialize inventory data, while HYDE used population density to spatialize data.
- 2) differences in statistical inventory data.
- 3) use of Boolean versus continuous description of the landscape, which makes direct comparison difficult.

## 4.3 The Land Use Scanner and scenarios for the year 2020

by H. Visser and J.A.M. Borsboom–van Beurden

In the following we will give an example of nominal map comparison for maps produced by the Land Use Scanner. The description of the Land Use Scanner is taken from Borsboom - van Beurden *et al.* (2002) and Scholten *et al.* (2001).

### 4.3.1 The Land Use Scanner

To support planning systems and evaluate the consequences of spatial development such as urbanisation, for example, it was decided in 1996 to jointly develop an instrument for the prediction of future land-use patterns for valuable landscapes, nature areas, the environment and water systems. Co-operating institutes are the RIVM, the Free University (VU), the National Spatial Planning Agency (RPD) and the Agricultural Economics Research Institute (LEI). The aim of the Land Use Scanner is to elaborate different scenarios by integrating and allocating exogenous land-use claims, coming from sectoral models such as housing and employment models. The resulting future land-use configurations may serve in turn as input for other models, such as hydrological and ecological models. The Land Use Scanner could serve in this way as an important mode of communication during a planning process (Schotten *et al.*, 2001).

At present the Land Use Scanner has been applied in several research projects. Here, the angle of research is not purely fundamental; in fact, it is mainly policy-oriented. An example of an application is the evaluation of various alternatives for the expansion of Amsterdam Airport (Schiphol). These include further concentration on the current location, relocation of Amsterdam Airport or the relocation of the runways to a new location in the North Sea, accompanied by a rapid transit connection for the management of passenger and freight flows on Amsterdam Airport. These alternatives were assessed in terms of the effects on the 'locational' patterns of employment and population. Another example is the simulation of Spatial Perspectives 2030. These perspectives, stated by the National Spatial Planning Agency in more-or-less qualitative terms, incorporate a wide diversity of assumptions with respect to trends in lifestyles, locational choices of households and companies, modal split and the future spatial organisation of the Netherlands.

The most extensive application of the Land Use Scanner hitherto has been the ex-ante evaluation of future land-use patterns in the preparatory studies of the Fifth National Policy Document on Spatial Planning (Goetgeluk *et al.*, 2000; MNP, 2001).

### 4.3.2 About the model

The Land Use Scanner has to be placed in the tradition of equilibrium models that use micro-economic-theory-based assumptions about supply, demand and price-setting (Schotten *et al.*, 2001). At the moment, the Land Use Scanner is predominantly used for the elaboration of scenarios, since it is not yet really suited for spatial optimization purposes. This means that more or less qualitative scenarios indicating future socio-economic developments serve as input to sectoral models to predict the future demand for land per distinguished land-use type.

Alternatively, expert judgement may be used to generate land-use claims when it is not possible to use models. Normally, the spatial unit is defined as grid cells of 500 by 500 metres, but this size can be adjusted if necessary. The classification of different land-use types is also flexible. For most applications, a classification of 15 land-use categories has been used, namely residential, industrial, roads, railways, airports, pasture, corn, arable land, flower bulbs, orchards, cultivation under glass, other agriculture, forest, nature and water.

For the allocation of future land use, three main parts of the models with corresponding datasets are of importance (see **Figure 18**):

1. Current land use. The current land-use map is based on the Land Use Statistics 1996 dataset from Statistics Netherlands and a classified remote sensing image, namely LGN-3. It derives its utility from the fact that the location with respect to certain land-use types determines the attractiveness, and in this way, the suitability for these land-use types. Besides this, it is also used for calculating total acreage per land-use type in the base year and the initial values of the balancing factors.
2. Land-use claims per land-use type. As stated earlier on, future land-use claims are generated by sectoral models or expert judgement, while policy agreements may have reserved areas for the realisation of its aims.
3. Suitability maps. These maps have been created per land-use type. Suitability maps, just as policy and physical environment maps, may include distance relations as well. Distance-relation maps contain attractiveness values, which, comparable to gravity modelling, are calculated by determining the distance to (current) types of land use with the help of distance-decay functions.

The mathematical elaboration of the model uses a logit approach. The probability that a certain category of land use is realized in a certain grid cell, is determined by comparing the suitability of this cell for this type of land use with the summed suitability for all other types of land use (see Schotten *et al.* (2001), and Hilferink and Rietveld (2001) for a more extensive discussion).

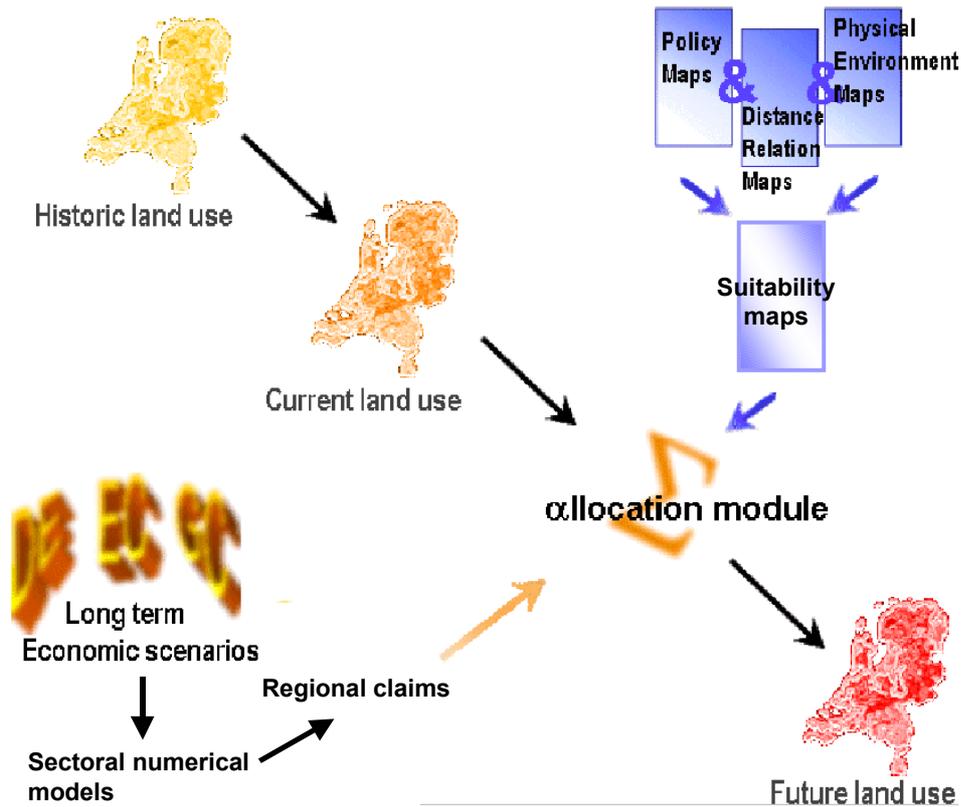


Figure 18 Flow chart of the Land Use Scanner model structure.

### 4.3.3 Scenarios

We will analyze two scenario simulations for the year 2020: the ‘*Country of Cities*’ scenario and the ‘*Country of Flows*’ scenario. According to the ‘*Country of Flows*’ policy concept, future land use will be directed to the optimal functioning of international streams, both in economic and ecological fields. This goal steers the choice of future locations for residential, work and nature areas. The economic ‘streams’ are directed to the road network and the ecological ‘streams’ to the water network. Combining of these two streams is the main challenge of this perspective. Urbanization will be concentrated along a few main transport axes in the form of a ‘string of beads’.

The ‘*Country of Cities*’ perspective foresees new residential areas with high intensities arising in and around existing cities. In this way the countryside will remain ‘open’. This openness will cause a new synthesis between nature, forest, water, agriculture and recreation to emerge. The borders between ‘living’ and ‘country’ will become sharper. Concentration of urbanization will increase the public support for collective transport.

More details on these scenarios can be found in Scholten et al. (2001). Both land-use scenarios are shown in the lower panels of **Figure 19** and the present-day situation (= 1995) in the upper panel.

#### 4.3.4 Comparing the different scenarios

To get a feeling for differences in present and future land use we have compared the present and projected land-use per category (compare section 2.1). **Figure 20** compares two categories: ‘agriculture’ (left panels) and ‘residential’ (right panels). Neither ‘agriculture’ panel shows much of a difference, while ‘residential’ panels show clear differences in the scenario perspectives.

For ‘*Country of Cities*’ we see a number of larger blue areas: ‘residential’ expands beyond the borders of ‘residential’ in 1995. Red cells (‘residential’ in 1995, not in 2020) are very small and randomly distributed over the area. However, for ‘*Country of Flows*’ we see new residential areas in the form of blue strings across the country (the ‘strings of beads’). Red areas are larger than those for ‘*Country of Cities*’.

But what is the actual difference between the two scenarios? Are future scenario differences due to minor differences in residential allocation? Or do these differences severely affect the way lands are used in the future?

In **Figure 21** shows the fuzzy-result map revealing differences between the two projections for the year 2020. The maps show the residential ‘strings’ in white, meaning that although there are differences between the two scenarios there is not a total difference (similarity around 0.50). However, the red cells in the result map indicate that a number of residential areas have been moved to the ‘string’ areas, while the ‘*Country of Cities*’ scenario had intensified these locations.

To be sure that the white and red differences in the figure are due to changes in residential (re-)allocation only, we calculated the result map for the corresponding two-category maps (‘residential’ versus ‘other’). We can aggregate categories within the MCK software by use of the Similarity Matrix in the Fuzzy set method (see section 3.1.2). The new fuzzy-result map (not shown here) is very similar to that shown in Figure 21. Thus, the most pregnant differences between the two projections are due to re-allocation of present residential to predefined string areas. This conclusion is in line with one of the ideas behind the ‘*Country of Flows*’ scenario: making the countryside more open.

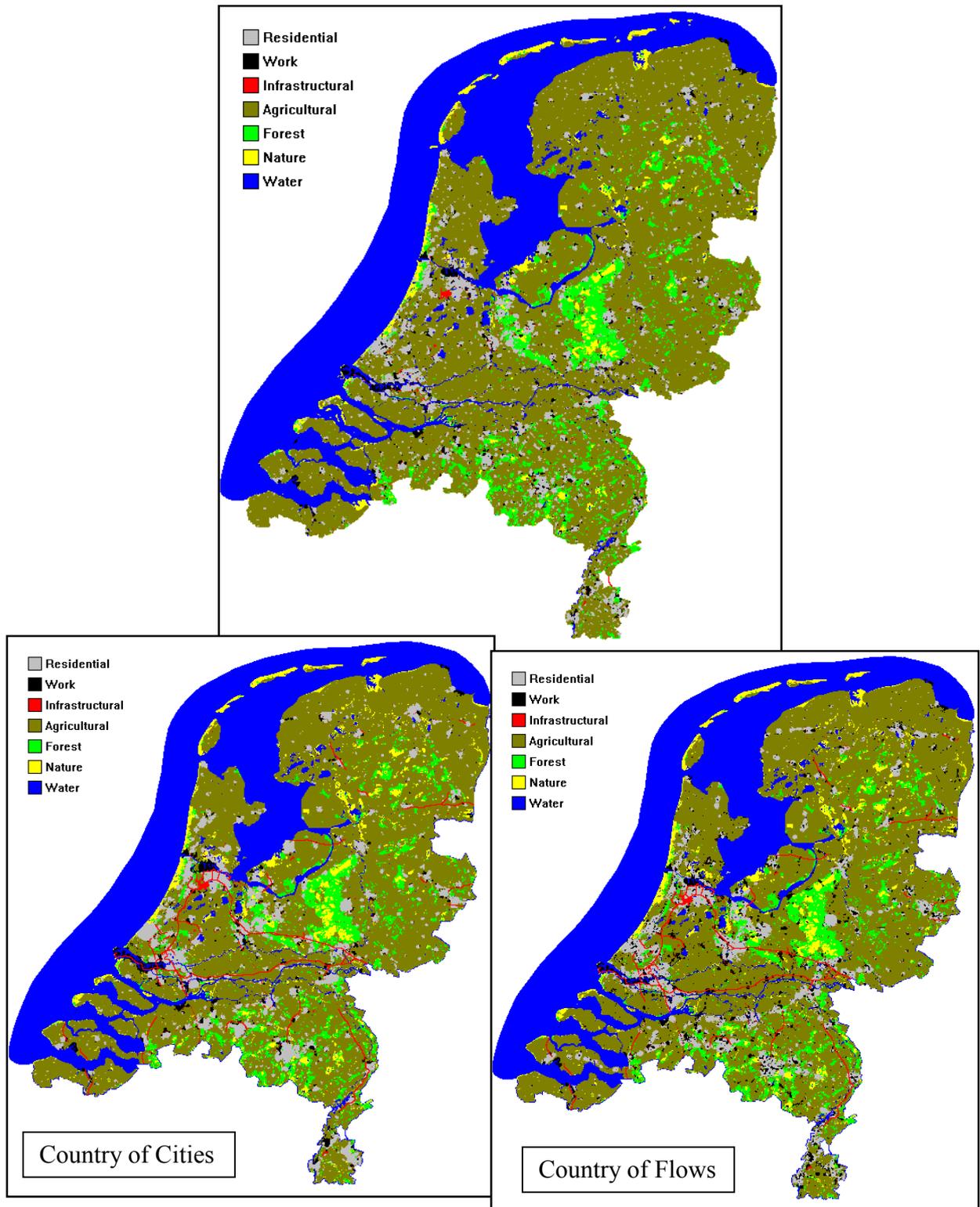
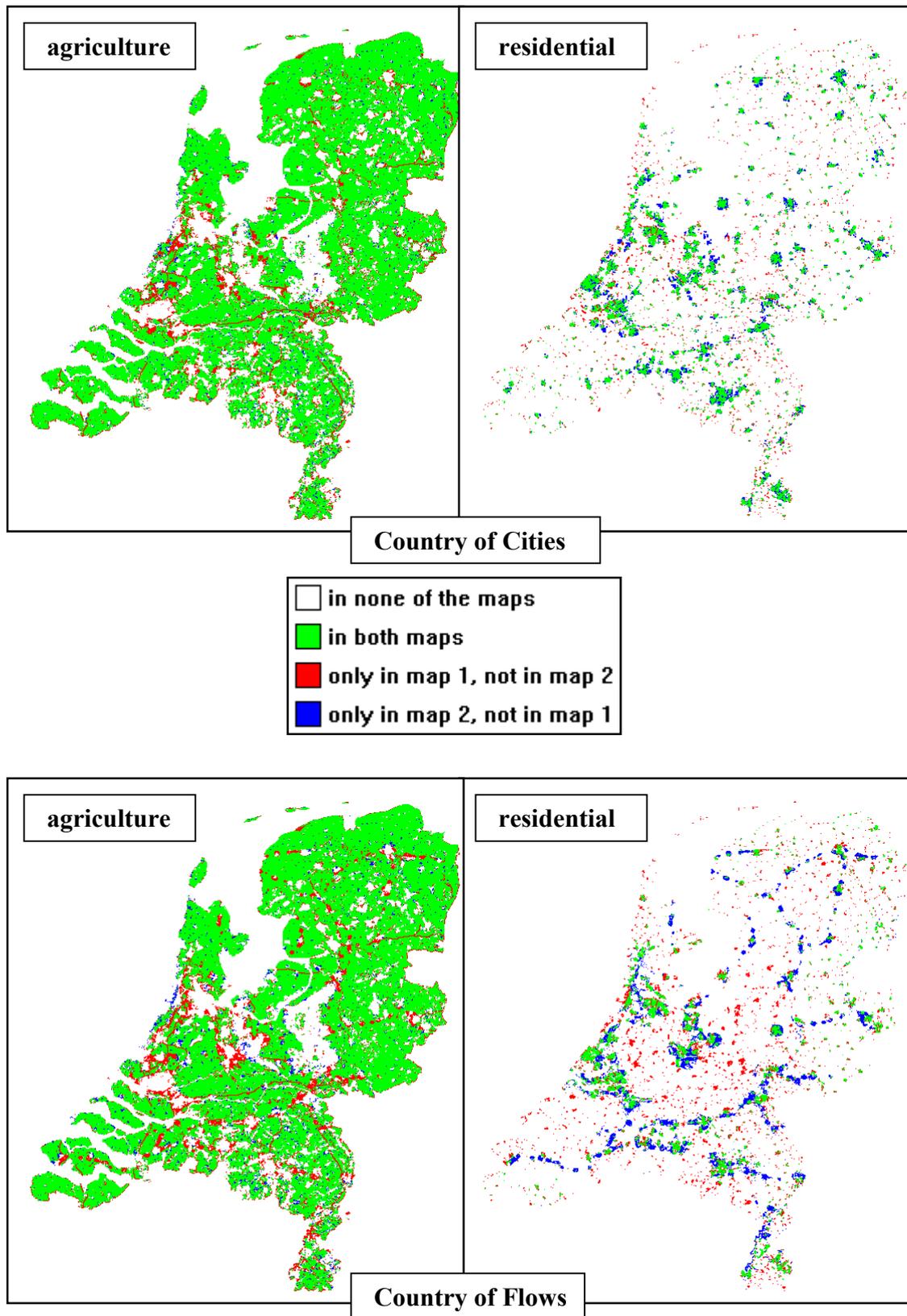
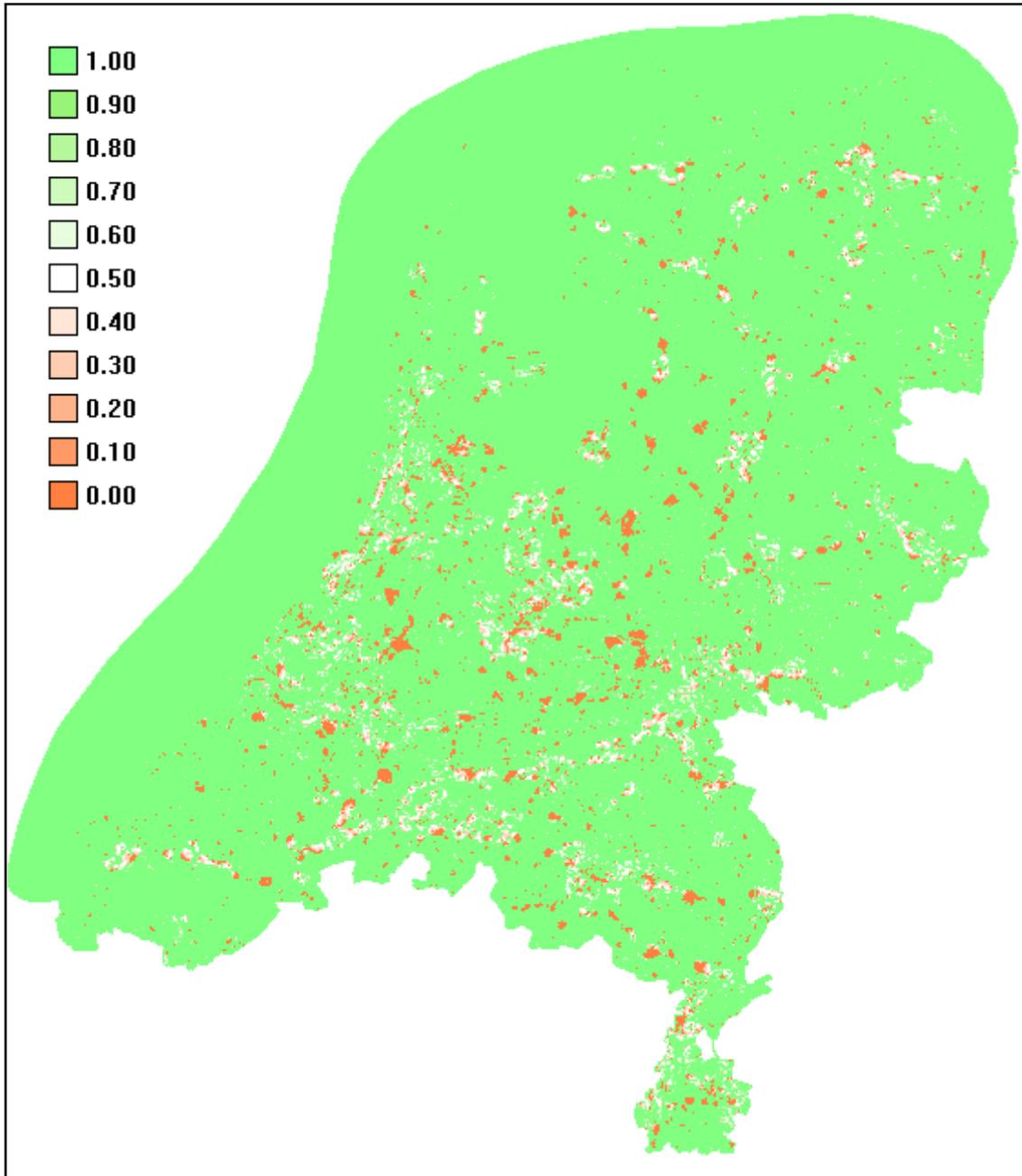


Figure 19 Land use for the year 1995 (upper panel), for the year 2020 according to 'Country of Cities' (lower left panel), and for the year 2020 according to 'Country of Flows' (lower right panel).



*Figure 20 Differences between the 2020 scenarios and the base map for 1995 (= map 1) per category.*

Upper panels are for the Cities scenario (= map 2) and lower panels for the Flows scenario (= map 2). Left panels are for agriculture and right panels for residential.



*Figure 21 Fuzzy result map for 'Country of Cities 2020' and 'Country of Flows 2020' (compare Figure 19).*

We have used the default settings from the fuzzy-set approach.

## 4.4 Residential land use in 2020

As examples of ordinal maps, we choose three maps from the Land Use Scanner for the whole of the Netherlands: residential in 1995, residential in 2020 for the ‘*Country of Cities*’ scenario, and residential in 2020 for the ‘*Country of Flows*’ scenario. The general characteristics of the Land Use Scanner were described in sections 4.3.2 and 4.3.3. The grid size is 500 by 500 m, and residential is presented in 11 percentage categories: no residential, 0 to 10% residential, 10 to 20% residential, etc. Clearly, the categories are ordinal.

The two maps for 2020 are shown in **Figure 22**. The projection for ‘*Country of Flows*’ clearly shows the openness of the countryside, as seen in section 4.3.4.

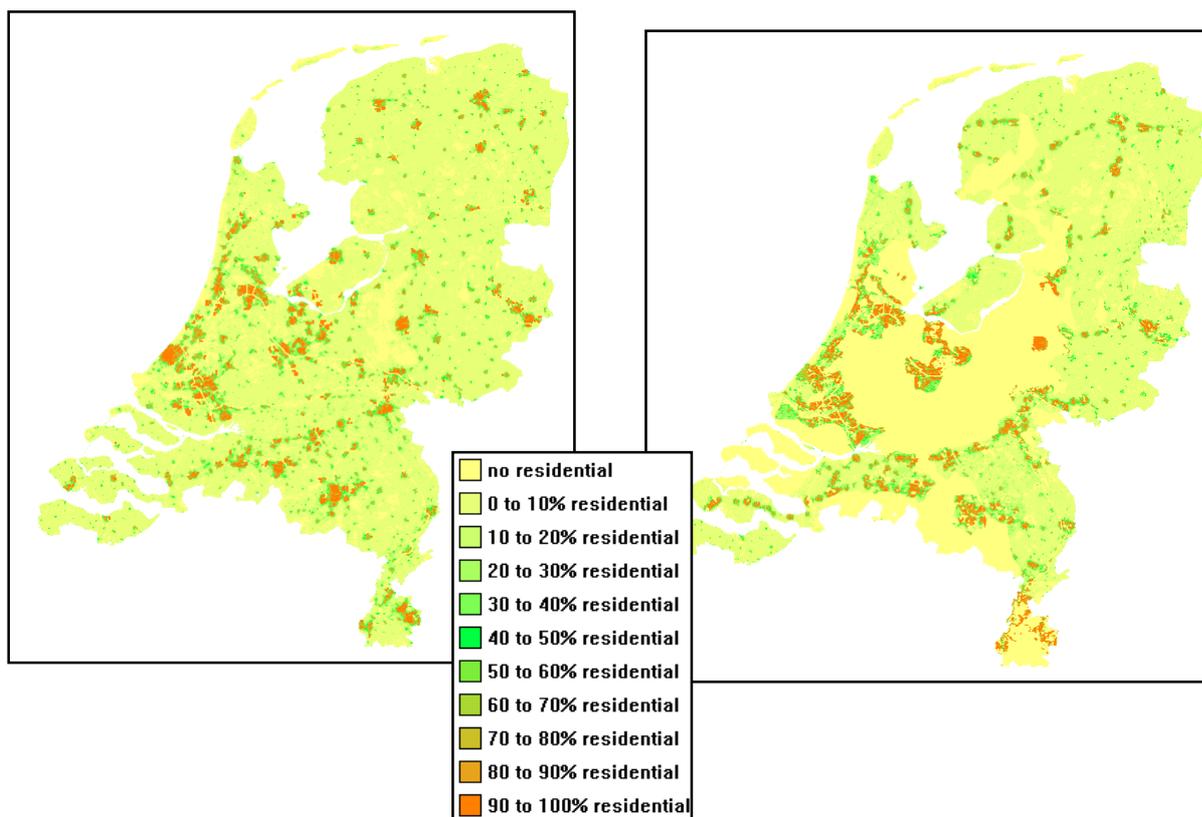


Figure 22 Percentage residential in 2020 for scenario ‘*Country of Cities*’ (left panel) and ‘*Country of Flows*’ (right panel).



*Residential simulation results by the Land Use Scanner are presented as a percentage of a particular 500-m by 500-m grid.*

*Photo: advertisement campaign Gemjobs.*

**Table 12** shows how the similarity between categories is defined. We have chosen values in the Category Similarity Matrix such that categories nearby (difference of 10% residential) have a similarity of 0.70. Categories differing by 20% residential have a similarity of 0.30. Other fuzzy settings are kept at their default values.

Using the settings from Table 12 we made two comparisons:

- observed residential 1995 versus residential ‘*Country of Cities 2020*’.
- observed residential 1995 versus residential ‘*Country of Flows 2020*’.

Table 12 Fuzzy settings and the Category Similarity Matrix for ordinal comparison of maps.

**Fuzzy Kappa Settings**

Settings Map 1

Radius of neighbourhood: 4

Exponential decay  
Halving distance: 2

Linear decay  
Slope: 0.5

Constant value  
Value: 0.5

Settings Map 2

Radius of neighbourhood: 4

Exponential decay  
Halving distance: 2

Linear decay  
Slope: 0.5

Constant value  
Value: 0.5

Buttons: OK, Cancel, Apply, Standard <<

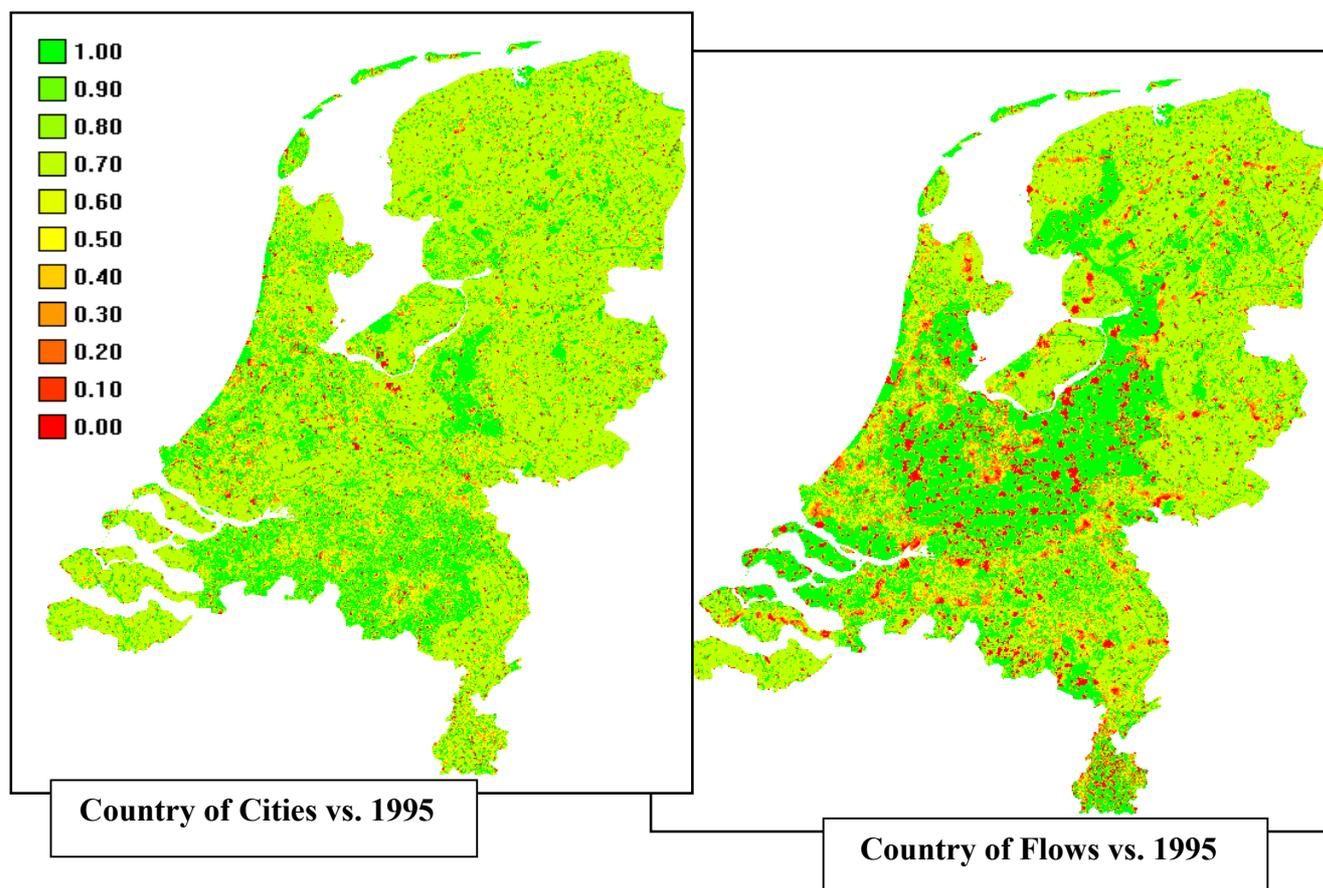
Category Similarity Matrix

Map 1 \ Map 2	no r...	0 to ...	10 t...	20 t...	30 t...	40 t...	50 t...	60 t...	70 t
no residential	1	0.7	0.3	0	0	0	0	0	0
0 to 10% residential	0.7	1	0.7	0.3	0	0	0	0	0
10 to 20% residential	0.3	0.7	1	0.7	0.3	0	0	0	0
20 to 30% residential	0	0.3	0.7	1	0.7	0.3	0	0	0
30 to 40% residential	0	0	0.3	0.7	1	0.7	0.3	0	0
40 to 50% residential	0	0	0	0.3	0.7	1	0.7	0.3	0
50 to 60% residential	0	0	0	0	0.3	0.7	1	0.7	0.3
60 to 70% residential	0	0	0	0	0	0.3	0.7	1	0.7
70 to 80% residential	0	0	0	0	0	0	0.3	0.7	1
80 to 90% residential	0	0	0	0	0	0	0	0.3	0.7
90 to 100% residential	0	0	0	0	0	0	0	0	0.3

Buttons: Load..., Save..., Unity

Both result maps are given in **Figure 23**. For ‘*Country of Cities*’ we see gradual differences between 1995 and 2020: much light green and yellow, not much dark green and a little red. For ‘*Country of Flows*’ we see the opposite: much dark green (identical grid-cell neighborhoods) and red (totally different grid-cell neighborhoods). Thus, changes in residential are much more striking in the ‘*Country of Flows*’ scenario.

This observation is also reflected in the Fuzzy Kappa values: for ‘*Country of Cities*’ we found Fuzzy Kappa = - 0.08, for ‘*Country of Flows*’, Fuzzy Kappa = - 0.87. We note here that negative Kappas imply similarities between maps which are *worse* than randomly produced maps containing the same number of cells for each land-use category.



*Figure 23 Fuzzy result maps contrasting Residential in 1995 and Residential in 2020 following the ‘Country of Cities’ scenario (left panel) and the ‘Country of Flows’ scenario (right panel).*

Maps are compared on an *ordinal* scale using the Similarity Matrix given in Table 12. For the left panel we found Fuzzy Kappa = -0.08 and for the right panel, Fuzzy Kappa = -0.87.

## 4.5 Travel times of ambulances

Three maps containing driving times of ambulances in the Netherlands are available in the MCK software. The maps, modelled by the Public Health division of RIVM (S.L.N. Zwakhals, VTV), show driving times on a 500 by 500 m grid expressed in minutes.

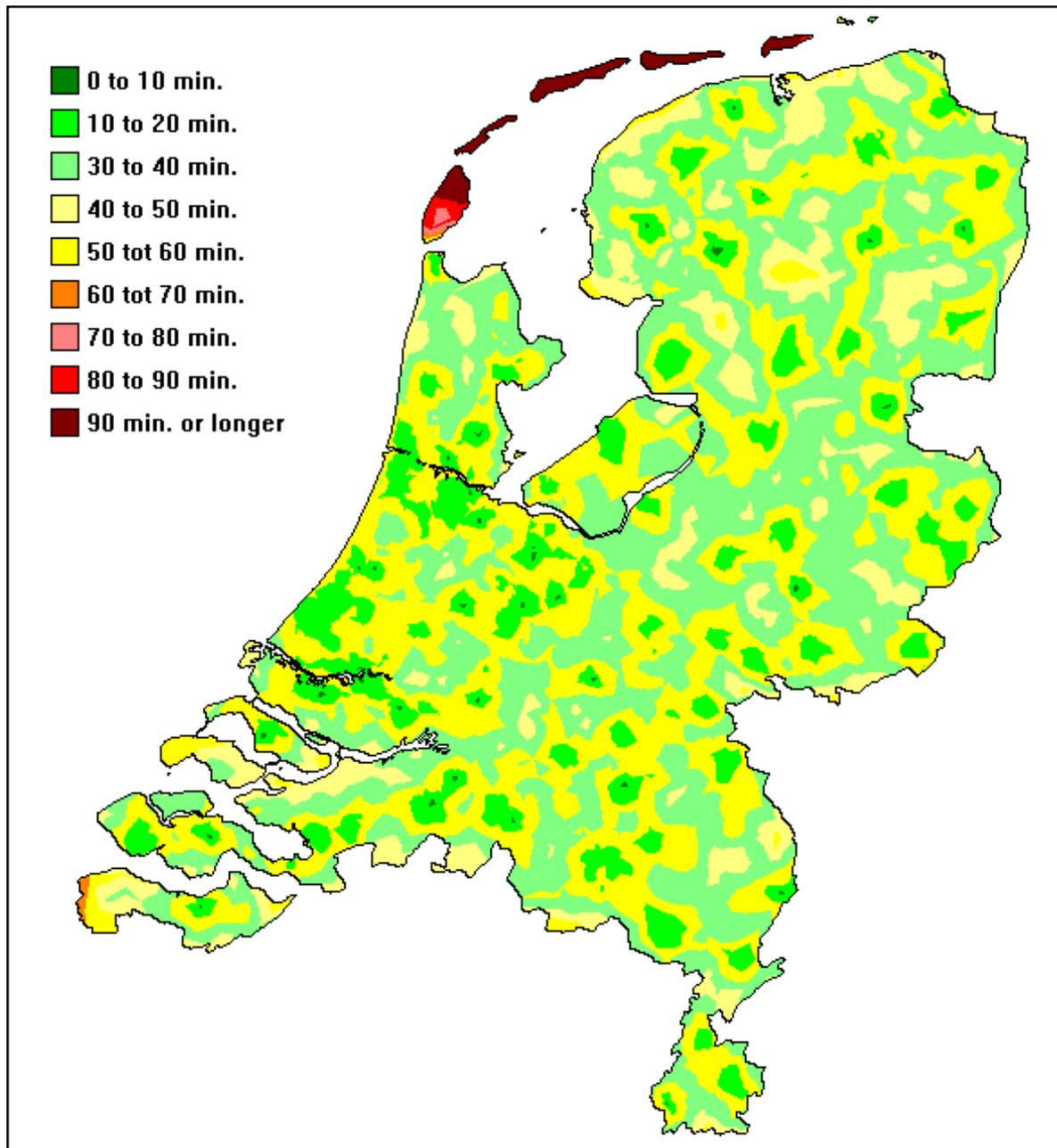
A driving time for a specific grid point is defined as the total time needed for an ambulance to get to that grid point and to bring the patient to the nearest Emergency Room (ER) hospital. Moreover, driving times have been augmented with 2 minutes for telephone response, and 5 minutes for the loading time of the patient. There are 200 ambulance garages in total and 110 hospitals with an ER in the Netherlands.

There are no governmental guidelines for these driving times. However, specialists call driving times up to 45 minutes 'ideal'. A driving time of 60 minutes, denoted as 'the golden hour', is known as a limit interval for an emergency admittance.

NB: Please note that there are governmental guidelines for driving times, defined as the time between the first call to the ambulance and the actual arrival at the patient's location. This driving time (response time plus driving time to the patient) has a limit value of 15 minutes. Maps showing these driving times can be found on the RIVM Internet site ([http://www.rivm.nl/vtv/data/atlas/zkvv\\_ambu\\_a016.htm](http://www.rivm.nl/vtv/data/atlas/zkvv_ambu_a016.htm)).

The MCK map, 'all110\_amb.asc', contains the driving time for the locations of 110 ER hospitals. The map 'all103\_amb.asc' contains driving times for 103 ER hospitals; in this map seven hospitals have been removed with respect to the first map. The map, 'sions\_amb.asc', contains driving times for 109 hospitals: here, only the ER hospital Talma Sionsberg in Dokkum was removed. Finally, we have a mask file to adjust driving times calculated by the model on a rectangular grid containing part of the North Sea and the IJsselmeer. Clearly, these grid points should not be incorporated in the analysis.

Driving time maps use a ratio scale. Furthermore, if we compare a map with 110 ER hospitals to one with 103 ER hospitals, we know that driving times will either be equal or prolonged (i.e. if hospitals are omitted). Therefore, there is no advantage in transforming both maps to an ordinal scale and analysing them by fuzzy-set methods (as was done in section 5.4). Simple subtraction or division of driving times for corresponding grid points will give informative result maps. Driving times for 110 ER hospitals are given in **Figure 24**.

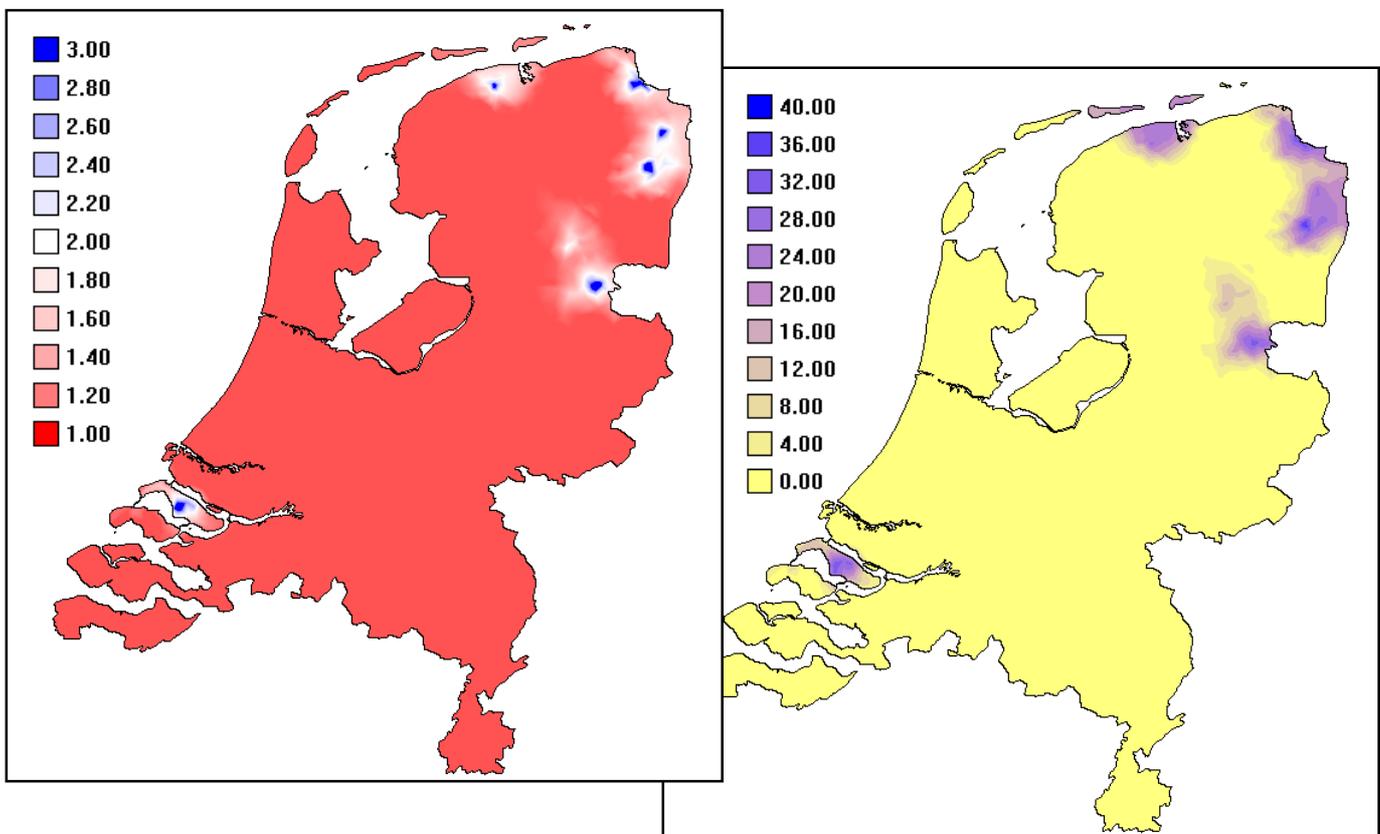


*Figure 24 Simulated ambulance driving times, given 110 ER hospitals and 200 ambulance stations.*

The driving times, expressed in minutes, represent the total sum of time required to response to the call, get to the location of the patient, loading the patient and then bring him/her to the nearest ER hospital. Driving times up to 60 minutes are preferred (the 'golden hour'). There are no specific governmental guidelines for these driving times.

In **Figure 25 (left panel)** we chose to divide driving times per grid cell (103 hospitals/110 hospitals). All values are found between 1.0 and 3.0. The locations of the seven hospitals omitted are easily seen in the figure. This panel also shows the area with white to blue (double to triple driving times) as being limited. The largest effects occur in the eastern part of the province of Groningen.

In **Figure 25 (right panel)** we subtracted driving times for corresponding grid points (103 hospitals – 110 hospitals). Here we can see that driving times are prolonged in some (small) areas by 36 to 40 minutes.



*Figure 25 Prolonged ambulance driving times due to the simulated closure of seven ER hospitals.*

The left panel is calculated according to the relation of:

$$\left[ \frac{\text{(driving times 103 hospitals)}}{\text{(driving times 110 hospitals)}} \right]$$

for each grid cell. The legend shows this relation in the range of 1.0 (minimum) up to 3.0 (maximum). The score 1.0 (colour red) means no change in driving times. In the right panel we used simple subtraction:

$$\left[ \text{(driving times for 103 hospitals)} - \text{(driving times for 110 hospitals)} \right] \text{ minutes}$$

The legend shows the difference starting at a minimum of 0 minutes up to a maximum of 40 minutes.

## 4.6 Heavy metals in topsoils

The presence of heavy metals in topsoils may lead to serious environmental problems. Although some heavy metals are essential for enzyme processes in living organisms (e.g. Cu and Zn), all heavy metals are harmful in relatively low concentrations. Different (groups of) organisms show clear differences in their sensitivity to the toxicity of these substances.

A choice was made for Cd, Cu and Zn in the construction of an index for toxic pressure by heavy metals in the Netherlands. The index formula, called the potential affected fraction of plant species (PAF), was given in Bakkenes *et al.* (2002). The PAF index was presented in Environmental Outlook 5 (MNP, 2000).



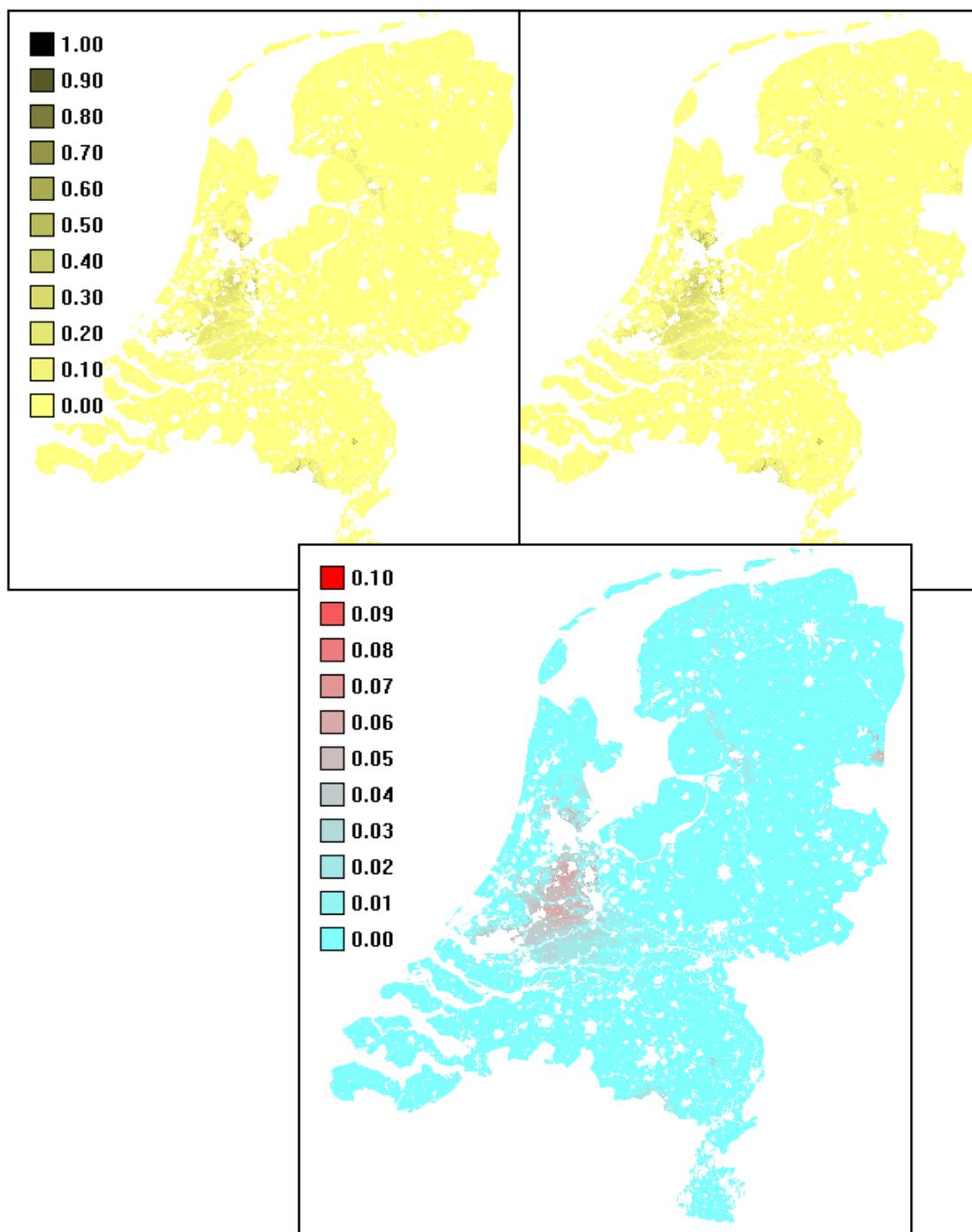
*Peat areas such as the Buurserzand are very sensitive to heavy metal soil pollution.*

*Photo: H. Visser.*

The upper panels of **Figure 26** show the PAF index for the years 1995 and 2030. For illustrative purposes we will show the results in distinctive classes of 10%. The most polluted area is between Amsterdam and Rotterdam. Maximum index values are around 0.70. Thus, extreme values near 1.0 (= deadly concentrations for plants) never occur. N.B.: to avoid misleading images, we did not put the legends in a range up to the maximum grid value of 0.70.

Inspection of both PAF maps showed that differences between the years 2030 and 1995 only showed constant index values or slight worsening over the years. This is why it is not very helpful to transform the maps to an ordinal scale (e.g. classes of 10%). A simple difference map will suffice here to show the PAF characteristics evolving over time.

A difference map is shown in the lower panel of Figure 26. Maximum differences occur around 0.10 and are located in the area south of Amsterdam.



*Figure 26* Upper left panel shows the PAF index for the Netherlands in the year 1995. The right panel shows the PAF for the year 2030.

The 0.0 value represents concentrations harmless to plants, while values of 1.0 would mean deadly concentrations for plants. The maximum grid value is 0.70. The lower panel shows the difference of the 2030 and 1995 PAF index values (subtraction of index values for each grid point). Differences are all positive, from 0.0 (no change) to a worsening of 0.10 (= 10%).



## 5. MAP COMPARISON KIT - special topics

### 5.1 Calibration of land-use models

by A.C.M. de Nijs and H. Visser

#### 5.1.1 Method

Land-use models have a number of ‘parameters’, not all of which are exactly known. In the case of cellular automata models such as the Environment Explorer, we have to choose a number of attraction decision rules or ‘distance-decay curves’ for specific categories. An example of such a curve is given in **Figure 27**.

Up to now these functions were chosen by visual comparison of maps. However, if we could quantify the difference between two land-use maps, one of which is a reference map, we could calibrate and even validate specific choices of decision rules. By calibration we arrive at more realistic decision rules; by validation we can judge the predictive power of the methods. For example, after calibration of the Environment Explorer to the 1996 CBS land-use map, we predicted land use for the year 2000 and compared the year 2000 simulation results to the 2000 CBS land-use map. Differences between maps can be quantified in terms of Kappas (sections 2.3.1 and 2.3.2) or Fuzzy Kappa (section 2.4.4).

In section 5.1.3 we will show how Kappa statistics can be used for the calibration functions of the Environment Explorer. The model itself will be described in short in the next section.

#### 5.1.2 Environment Explorer

The Environment Explorer is a spatial, dynamic model in which land use and the effects on social, economic and ecological indicators are modelled in an integrated way (Engelen *et al.*, 2003). Its primary goal is to explore future developments, combining autonomous developments with alternative policy options in relation to the quality of the environment where the Dutch citizens live, work and recreate. Various policy options from governmental departments are translated into a spatial, dynamic image of the Netherlands’ future with respect to issues such as: economic activity, employment, social well-being, transportation and accessibility, and the natural environment.

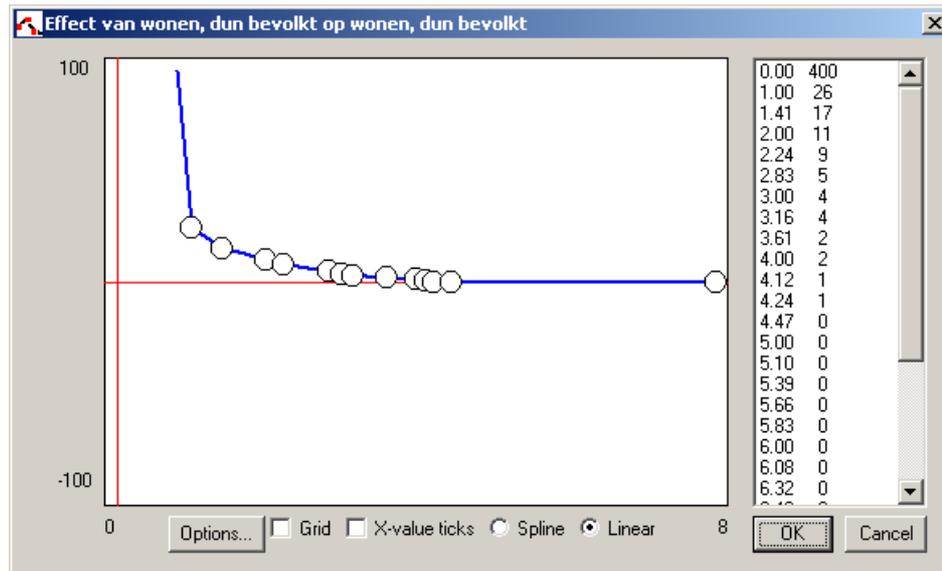


Figure 27 Example of a cellular automata distance-decay curve for the category 'sparse residential'.

The x-axis represents distances between cells with 'sparse residential' (data are shown at the right). Values above zero (the horizontal red line) point to attraction of 'sparse residential' cells at a particular distance and negative values to repulsion. Shown are the standard distance-decay curves within the Environment Explorer: positive attraction up to distances of 4.3 cells ( $\equiv$  2.15 km).

The Environment Explorer is characterized by three scale levels (see **Figure 28**):

- National level. The model combines national economic, demographic and environmental growth scenarios and passes them on to the regional level.
- Regional level. The Environment Explorer models the localization of new activities as a result of national and regional growth and arranges the relocation of activities and people in the 40 administrative COROP regions of the country. The resulting regional developments can be restricted to a range between a lower and upper maximum level per region or set to a fixed value.
- Local level. The Netherlands is represented as a grid of cells with each cell representing land use of a parcel of land of 500 by 500 m. A Constraint Cellular Automata model performs the dynamic allocation of the regional claims for space. The local level returns global information about the availability and quality of land for further expansion the COROP and national levels, thus influencing the dynamics at these levels.

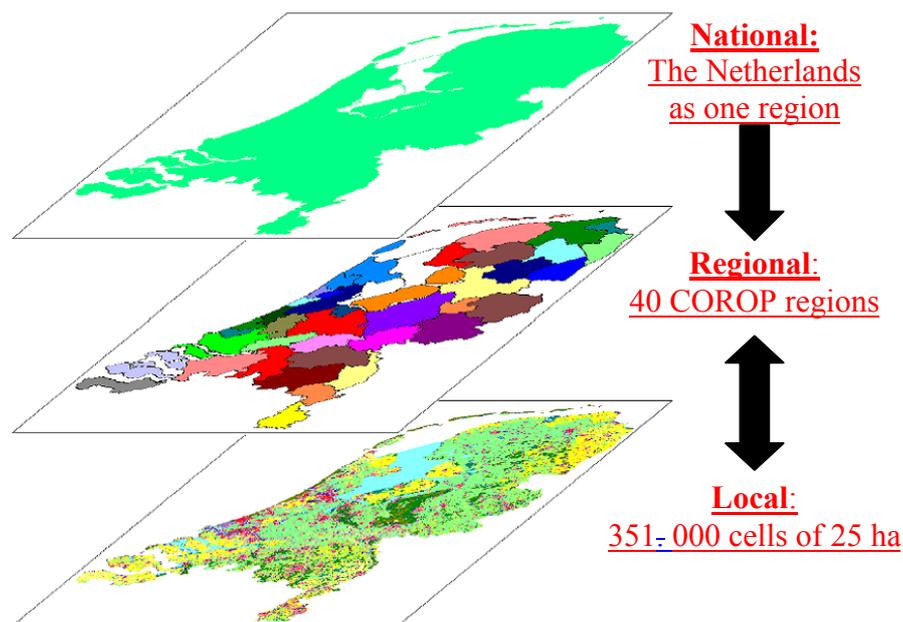


Figure 28 The three spatial levels represented in the Environment Explorer: national, regional and local.

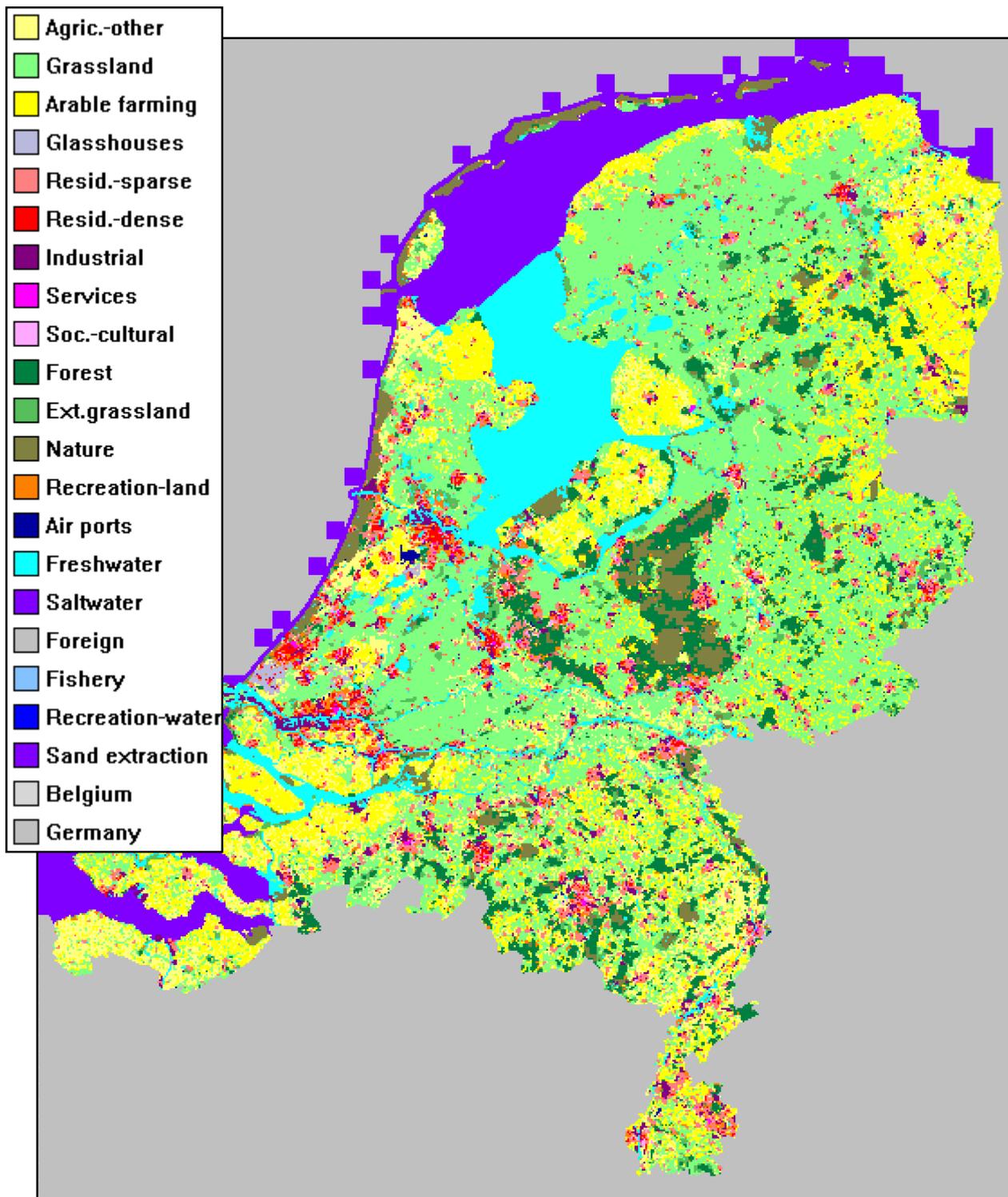
The transition potential (TP) determines allocation at the local level. The transition potential is defined as the weighted sum of the neighbourhood potential (NP), the policy or zoning map (ZM) and the suitability map (SM):

$$TP_{k,i} = NP_{k,i} \times (c_1 \times ZM_{k,i} + c_2 \times SM_{k,i})$$

in which  $c_1$  and  $c_2$  are two weighting factors;  $k$  is the land use and  $i$  the cell indices. The neighborhood potential is defined by the set of rules for the Constraint Cellular Automata. These rules describe the influence of all the land uses within a radius of 8 cells on the neighborhood potential. The relative attraction or repulsion for each land-use combination is described by a *distance-decay curve* (compare Figure 27).

The Environment Explorer translates national scenarios into dynamic maps on the regional and local level, showing land-use changes year after year. The driving forces behind these dynamics are the effects of economic and demographic developments on the land market: demand and supply, and the relative power of the various land users to realize their claims for space. For a detailed discussion including the mathematical description, the user is referred to De Nijs *et al.* (2001).

An example of a standard simulation for the year 1996 is given in **Figure 29**.



*Figure 29* Land-use predictions for the year 1996.

The Environment Explorer takes the actual situation in 1989 as the starting point.

### 5.1.3 Calibration of the Environment Explorer for 1996

By varying the form of a distance-decay curve, such as shown in Figure 27, we can select the optimum curve by comparing simulated 1996 maps to reality, i.e. the 1996 CBS land-use map. An alternative curve is given in **Figure 30**.

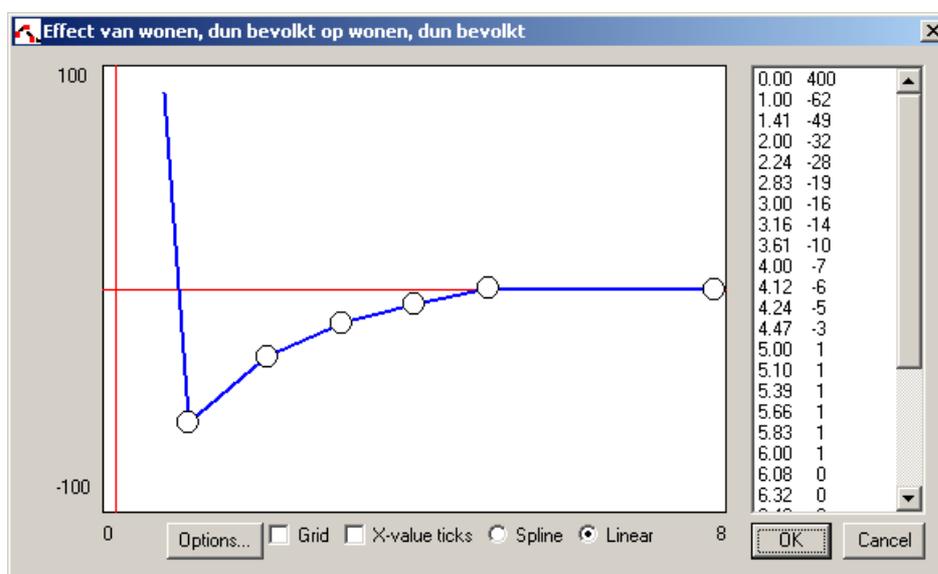


Figure 30 Alternative distance-decay curve.

**Table 13** shows four Kappa statistics. The first column compares the Environment Explorer simulation using the standard distance-decay curve for ‘sparse residential’ (Figure 27). The second column the same but now based on Figure 30. The table shows the standard distance-decay curve to perform (somewhat) better.

The third column shows the Kappas between the CBS 1989 map and the CBS 1996 map. Here, we used the 1989 map as a persistence predictor for the year 1996 (i.e. nothing changed between 1989 and 1996). It is interesting to note that the persistence predictor scores best in all four Kappa statistics. This result is in line with a claim of Gill Pontius that persistence predictors always outperform simulations by land-use models (it is a challenge for modellers to disprove his claim; a counter example was given for the Dublin case, Figure 9).

Table 13 *Kappa statistics.*

	<b>Map comparisons</b>		
	First simulation 1996 versus CBS-1996	Second simulation 1996 versus CBS-1996	CBS-1989  versus CBS-1996
<b>Kappa</b>	0.89	0.87	0.92
<b>K-location</b>	0.91	0.91	0.94
<b>K-histo</b>	0.98	0.96	0.98
<b>Fuzzy Kappa</b>	0.88	0.85	0.91

In Table 13 Kappas are shown on the basis of all 17 land-use categories. If we are only interested in the performance for ‘sparse residential’, we could calculate Kappas for maps containing two categories (‘sparse residential’ and ‘other’). In this way the calibration process is more directly tuned to the category under consideration (i.e. ‘sparse residential’). Kappas are shown in the first three columns of **Table 14**. As a reference we also calculated Kappas for ‘dense residential’ (three columns from the right in the table).

Table 14 *Kappa statistics for two-category maps.*

	<b>Sparse residential</b>			<b>Dense residential</b>		
	First sim. 1996 versus CBS-1996	Second sim. 1996 versus CBS-1996	CBS-1989  versus CBS-1996	First sim. 1996 versus CBS-1996	Second sim. 1996 versus CBS-1996	CBS-1989  versus CBS-1996
<b>Kappa</b>	0.90	0.71	0.89	0.91	0.87	0.94
<b>K-location</b>	0.91	0.84	0.94	0.92	0.87	0.96
<b>K-histo</b>	0.98	0.85	0.95	0.99	1.00	0.98

For 'sparse residential' the table shows that the highest Kappas are found for the first simulation (using the standard distance-decay curve). Here, Kappas are even better than the persistence predictor! For 'dense residential' we found the typical result of a slightly better estimate for the first simulation than for the second simulation. The persistence predictor outperforms both.



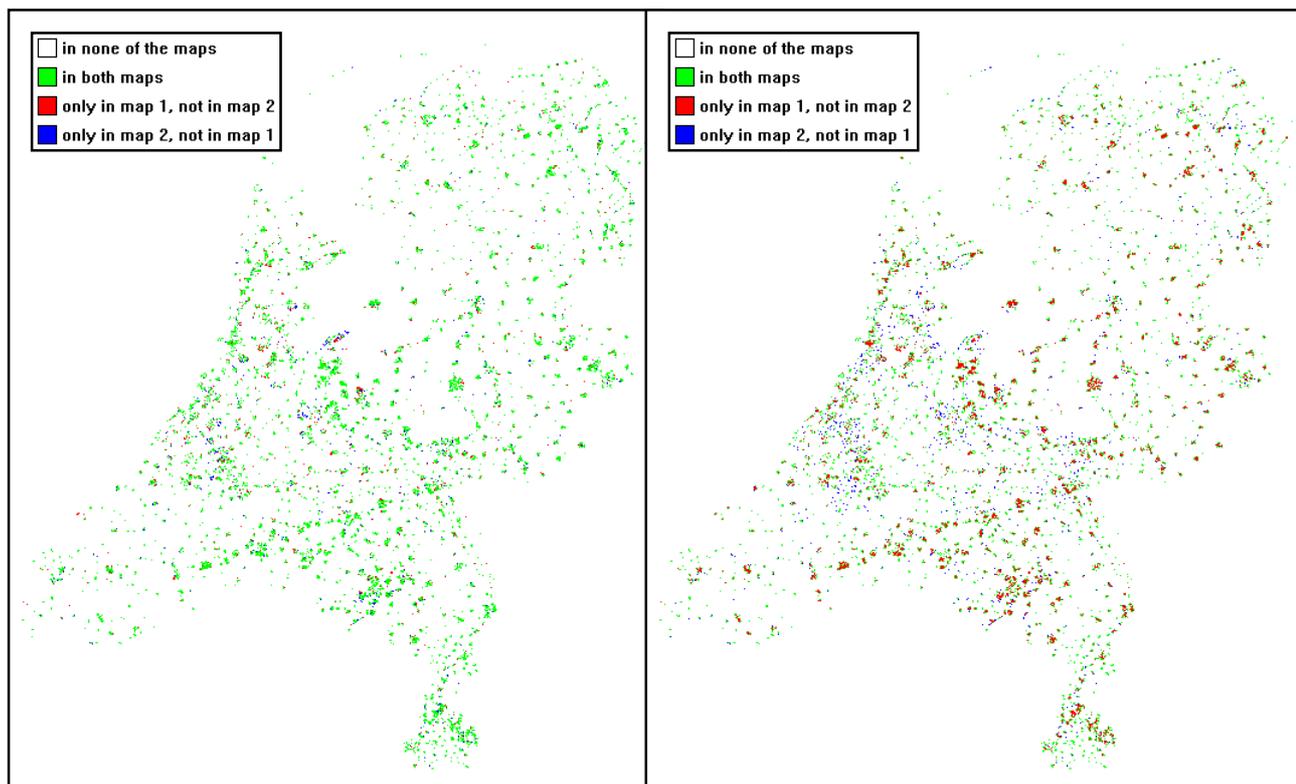
*Prediction of land-use changes by simple extrapolations of the present situation will not give satisfactory results in the long run, as illustrated by the temporal sequence of photos. These cyclorama photos were taken from April 1999 to June 2002, near Terneuzen. The lands of the farm 'Hoeve Nooit Gedacht' were 'transformed' to a traffic circle near the Westerschelde tunnel. It is a challenge for land-use models to predict these types of infrastructural transformations.*

*Photos: S. Swart.*

In addition to the first two columns of Table 13, we calculated the result maps per category (compare section 2.2 for an explanation). In **Figure 31** the left panel shows the first simulation and the right panel, the second simulation. A great number of predicted cells with ‘sparse residential’ are wrongly projected (colour red in the right panel) due to the unrealistic distance-decay curve for the second simulation.

We conclude that Kappas, and in particular, Fuzzy Kappa are very well-suited for calibration and validation purposes. This is especially true if we want to mimic human judgement. If we vary parameters corresponding to one specific category, we could choose to calculate the Kappa values for the two-category maps (‘specific category’ versus ‘other’).

Currently, an adapted version of the MCK software is in development by RIKS for semi-automatic calibration of the Environment Explorer.



*Figure 31 The CBS-1996 map (map 1) and the Environment Explorer simulation for 1996 (map 2) are compared for the category ‘sparse residential’.*

In the left panel we applied the default distance-decay curve for sparse residential (Figure 27) and in the right panel, simulation of a deviating (wrong) setting for ‘sparse residential’ (Figure 30).

## 5.2 Consistency in land-use maps

by R. de Niet and A.C.M. de Nijs

This section describes an application of the MAP COMPARISON KIT in the process of designing a consistent set of land-use maps for spatial models in the context of the LUMOS toolbox. The LUMOS toolbox consists of the Land Use Scanner (sections 4.3 and 4.4) and the Environment Explorer (section 5.1). Due to errors and definition changes, the original CBS land-use maps are less suited for temporal trend analysis and calibration purposes. We have used the MCK software to find out what land-use changes are *real* and which ones are not. Using these findings we have been able to make a set of correction rules to adapt the original CBS land-use maps. The procedure has yielded a more consistent set of land-use maps.

### 5.2.1 Context

Spatial land-use models such as the Environment Explorer (LOV in Dutch) and the Land Use Scanner make use of a base set of land-use maps. On the one, these maps are used for trend analysis and calibration of the model. On the other, the maps are the starting point from which the model allocates new land uses up to 2030 or 2050. The land-use maps in the Environment Explorer and Land Use Scanner are based on the ‘Bodemstatistiek’ land-use maps for the years 1989, 1993 and 1996 (CBS, 1993; CBS, 1996; CBS,1999). These maps are aggregated and re-sampled to a coarser grid size before we can use them in the models (section 5.2.2).

When we first explored these maps in the process of calibration and trend analysis, we noticed some suspect changes over the 1989-1996 period: here and there housing or industrial areas disappear, nature areas change to water, airports shrink in size and forests disappear. We used the MCK software to find out to what extent and where these unexpected changes appear. On the basis of these findings we defined a set of correction rules to make changes to the original CBS land-use maps (section 5.2.6). In this way we produced a more consistent set of land-use maps.

In the following sections we will describe the correction process with special attention paid to the role of the MCK software. This aspect comes up in sections 5.2.3 and 5.2.4.

### 5.2.2 From CBS land-use maps to LUMOS land-use maps

The CBS land-use maps are made by digitizing and interpreting aerial photography. The maps identify 35 different types of land uses in grid cells of 25 m. For use in the LUMOS toolbox, and more specifically in the Environmental Explorer model, we aggregate these land uses to 17 types and resample the grid to a coarser grid size of 500 m. For the re-sampling process we use the so-called SPAT-tool (Meulen, 1999). The SPAT tool (SPatial Aggregation Tool) allows us to make it possible to keep the total level of each land use the same between the 25-m and the 500-m land-use map.

In 2001 RIVM made a detailed study of possible errors in the CBS maps when using them to identify land-use changes over time (Raziei and Evers, 2001). This study revealed that errors or unexpected changes were the result of:

1. improved techniques for the interpretation of aerial photography and more detail in photographs between the first year (1989) and the last year (1996).
2. minor changes in the definition of land-use categories between 1989 and 1996. For instance, the grass along highways was identified as 'road' in 1989 and as 'other uses' in 1996.
3. a shift from describing 'land cover' instead of 'land use'. For instance, the grass surrounding a hospital was formerly seen as part of the hospital (category: 'socio-cultural provisions'), but later classified as parks and public gardens.
4. (real) errors in the digitizing process.

In 2001 category 4 errors, e.g. more or less detail in infrastructure, were corrected and made consistent in the 1989, 1993 and 1996 maps. These maps serve as input for the LUMOS set of land-use maps.

### 5.2.3 Locating and identifying suspect changes by the MCK

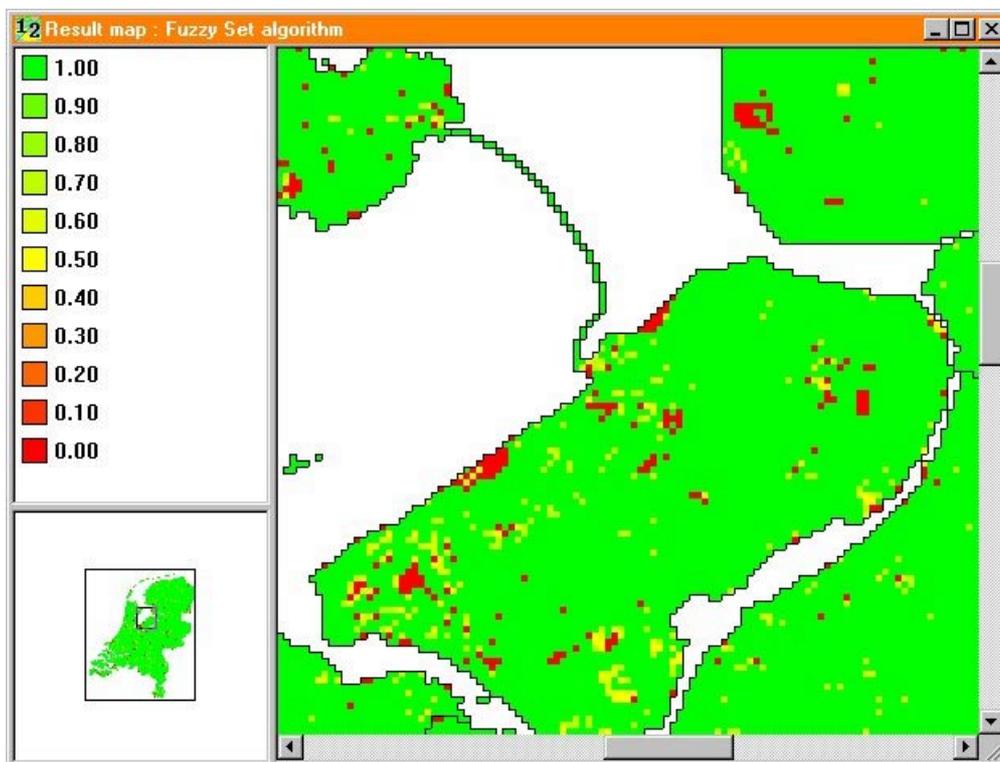
In the process of identifying suspect changes, an important step is to answer the question of 'what are suspect changes? We decided to investigate the following changes on a common-sense basis:

- nature that disappears
- built-up area (residential or commercial) that disappears
- shrinking airports
- water (salt or fresh) that disappears
- recreational areas that disappear
- glasshouses that disappear
- changes from one type of nature to the other (forest, extensive agricultural use and nature reserves)
- an increase in the size foreign countries (Belgium and/or Germany).



For our purpose we grouped the original categories into four larger groups ('housing', 'work', 'nature' and 'agriculture') and a few smaller groups. The value for the 'radius of neighbourhood' is crucial in this comparison, because it can disguise changes if not set properly. It was recommended to vary the value between 0 and 4 and look for the best result.

**Figure 32** shows the result map for the fuzzy-set algorithm comparison between the 1989 and 1996 land-use maps. Using the fuzzy-set algorithm we found the largest changes to be between the 1989 and the 1993/1996 land-use maps. It also revealed that the differences between the 1993 and the 1996 maps were smaller. This is in accordance with the documentation of Raziei and Evers (2001). A simple explanation for this is found in the 1989 map, the first in a range of maps made using a new technique for making land-use maps called digitizing land use by use of aerial photography.



*Figure 32 The result map with the fuzzy-set algorithm comparing the 1989 and 1996 land-use maps for the Flevoland area.*

### 5.2.5 Categories in detail

On the basis of the findings from the Fuzzy-comparison method we could selectively use the categorical comparison method to answer the following questions: what is the change' and 'is the change suspect or not'? The next two examples illustrate the changes we found for airports and nature areas.

The first example shows the area of Schiphol, the Netherlands main national airport. **Figure 33** clearly shows the area occupied by airports to decrease between 1989 and 1996. This is an unexpected change since we know that the built-up area of the Schiphol region only *expanded* during this period. Looking at the land-use maps for 1989 and 1996 (lower panel in Figure 33), we see that part of the airport has been put to a (new?) industrial use. It is likely that these areas were either grassland or already in use as industrial areas in 1989. So this change is due to two effects: distinguishing more detail in the 1996 land-use map and the shift from describing 'land cover' rather than 'land use'.



*Inconsistencies in categories across land-use maps may arise for a number of reasons. Several category definitions may apply to the grid shown here, e.g. are we dealing with 'water', 'industrial', 'residential', 'recreational' or 'other agricultural'?*

*Photo: H. Visser.*

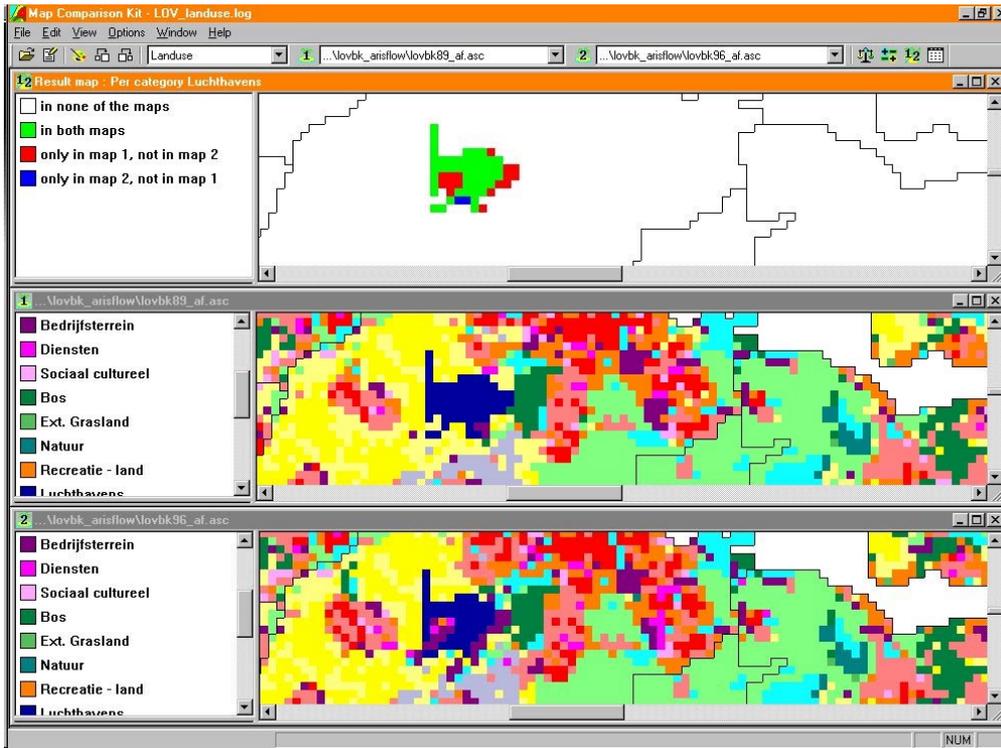


Figure 33 Land-use change for 'airports', i.e. the Schiphol area.

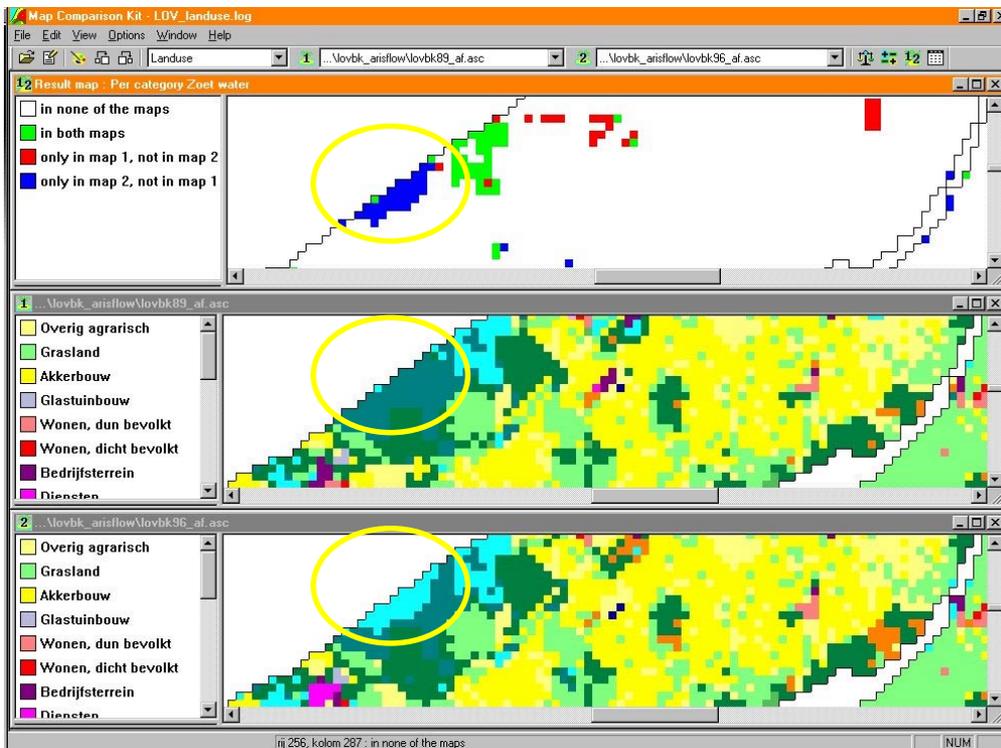


Figure 34 Land-use change for 'water', the area of the 'Oostvaardersplassen' in the Flevo Polder.

The second example shows the area of the Oostvaardersplassen, a large nature reserve in the Flevo Polder (**Figure 34**). When looking at the category ‘fresh water’, we notice a large increase over 1989-1996 period. A look at the land-use maps for 1989 and 1996 (lower panel in Figure 34) shows that nature in 1989 becomes water in 1996. This is an unexpected change: a large part of the Oostvaardersplassen does, in fact, consist of water, but the whole area is still designated as a ‘nature reserve’. So this change is a clear example of the shift from describing ‘land cover’ rather than ‘land use’. In a range of land-use maps we do not want these kind of changes: ‘nature’ has to stay ‘nature’.

Looking at land-use changes in this way, one has to keep in mind that a change may look suspect, but can be real. For instance, in the case of the housing project, IJburg near Amsterdam, ‘water’ becomes ‘housing’. Another example is the case of industrial use (extraction of raw materials) changing to nature and water. This seems strange but it is a real change: after the excavation of peat diminished, the area was transformed into a nature reserve.

From all the possible suspect changes identified, we finally decided to devise a set of correction rules for five of them. The other did not appear or appeared, but only on a small scale. These changes are:

- ‘nature areas’: takes on a ‘water’ or ‘agricultural use’.
- shrinking of airports.
- ‘water’ that disappears and takes on ‘agricultural use’.
- industrial areas that disappear between 1989 and 1993, then reappear in 1996.
- Changes from one type of nature to the other (forest, extensive agricultural use and real nature).

## 5.2.6 Formulating correction rules

When making final map corrections, we chose application to the CBS land-use maps and not the LOV land-use maps. In this way, other researchers may use the corrected version for their own applications. Using the located suspect changes, we can selectively look at the more detailed CBS land-use maps (25 m vs. 500 m, and 35 vs. 17 categories; see also section 5.2.2) to find out which land-use category is under. This was done using standard GIS software (Arcinfo 7).

After careful inspection we formulated seven correction rules (or rather ‘if...then’ statements) for the 1989 map and three for the 1993 map. Through these correction rules we made a comparison among the three CBS land-use maps and then applied the inferred changes to the 25 by 25 m CBS maps.

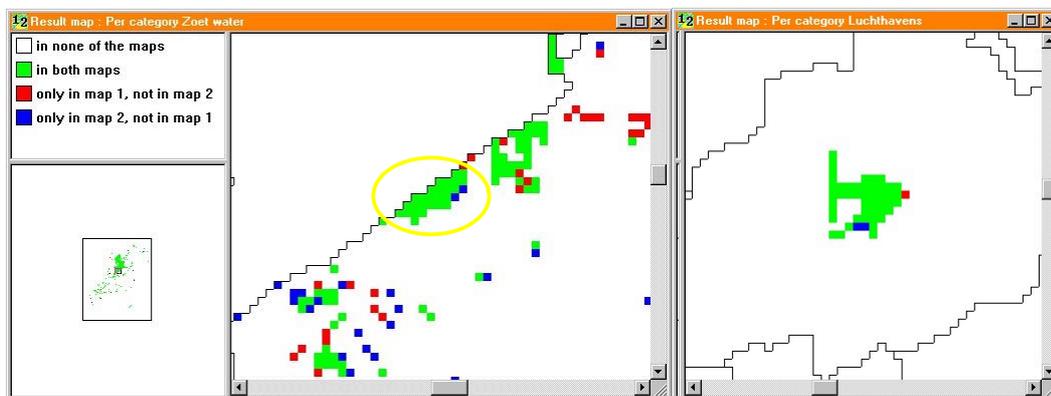
Two examples of correction rules for the 1989 map are:

- if the land use in 1989 is water but not in 1993 and 1996, change the land use to the land use of 1993
- if the land use in 1989 is airport but not in 1993, change the land use to the land use of 1993

An example of a correction rule for the 1993 map is:

- if the land use in 1993 is not nature, but is nature in 1989 and 1996, change the land use to the land use of 1996.

On the basis of the corrected set of CBS land-use maps we generated a new set of 500 by 500 m land-use maps for the Environment Explorer (see also section 6.2.2). **Figure 35** shows some of these maps for the Oostvaardersplassen and the Schiphol region.



*Figure 35 Land-use changes in 1989 and 1996 for fresh water (left panel) and airports (right panel) for the regions Oostvaardersplassen and Schiphol, respectively. Both panels are based on the new 500 m by 500 m LOV land-use maps derived from the corrected version of the CBS land-use maps.*

The new maps are more consistent and better suited for calibration and temporal trend analysis purposes. However, not all inconsistencies can be solved in this way. Thus, one still has to be careful in interpreting land-use changes.

### **5.2.7 Conclusions**

The MCK is a helpful instrument in the process of making distinctions between real land-use changes and errors in series of land-use maps. The MCK makes things easier and gets results faster. On the basis of the outcomes, we were able to formulate correction rules for the CBS land-use maps. This has led to a more consistent set of land-use maps better suited for the detection of trends over time and for calibration purposes (compare section 5.1).



## 6. Conclusions and further research

This report has summarized map-comparison methods in which both traditional and more advanced techniques have been discussed. We have paid more attention to the treatment of nominal maps since existing methods are sparse here. However, handling maps on ordinal and interval/ratio scales has also been discussed.

The applications in Chapters 4 and 5 have made clear that visual inspection of maps, cell-by-cell comparison, and fuzzy-set map comparison are able to grasp the characteristics of map differences, although in different ways. The examples have shown that the three approaches combined will be an ideal way of analyzing maps and map differences. In fact, we are combining ‘the best of three worlds’.

All examples, except the one in Appendix A, have been performed using the MAP COMPARISON KIT software. Our experiences are very positive. The software is easy-to-use and there are no limitations to the size of the maps (limits are set by the memory conditions of the PC). The comparison techniques are state of the art. Still we have some comments and additional wishes.

Hereafter, we will describe our comments/wishes in three categories:

1. User-friendliness of the software
2. Additional features
3. Integration of the software into other GIS environments.

### Ad 1 Use of the software

- Software has no help function. Although this ‘gap’ is filled by a good user’s guide, a help button would ease the use of the software for inexperienced users.
- It would be of help if the software could produce a *meta data file* when exporting maps. These meta data should include details on how the map is generated.
- All computation results must be saved by hand. It would be helpful if results could be saved automatically in a ‘history file’.
- If maps to be compared contain ‘long borders’, the calculation of the different Kappa statistics is not optimal. With ‘borders’ we mean all transitions from observed cells to cells with missing values. The Kappa’s should be corrected for ‘edge effects’.

- It would be helpful if several categories could be re-classified into one new category within the MCK software. E.g., it could be of help to re-classify the categories ‘nature’, ‘forest’ and ‘grassland’ to a new category ‘green space’. Currently, such a re-classification must be performed outside the software, for example, through the MAPTRANSFORMER software (see section 3.2.1). Alternatively, we could use the advanced screen of the Fuzzy-set map comparison method. However, in this case, we do not have the possibility to calculate the various Kappa values. An ‘on-the-fly’ re-classification, performed on a predefined set of maps, would be of great help.
- If a mask file, also denoted as region map, is used within the MCK and if we want to copy a masked map to Word, the map is copied **without** the mask. This is not what we want in most cases.
- Changes in the region map should be performed *outside* the software. It would be of help to have flexibility in the region map *within* the software. E.g., it would be of help if we could calculate Kappa statistics for individual provinces within the Netherlands, having data for the whole of the Netherlands. In the present version of software we should provide region maps for each individual province.
- It would be of great practical importance to have the option to zoom in on a certain rectangular or circular area in one of the two maps selected for comparison. Now, we should be able to calculate map statistics only for this *selected* area in both maps. Then, we can compare these local estimates with the overall (Kappa) statistics. In this way the stability of (Kappa) statistics can be tested interactively. Also the detection of hot-spots in one of the two maps will become easy by this zoom option.
- It would be helpful to have a batch environment around the MCK software for handling huge amounts of maps. In some NEAA applications over 100 pairs of maps should be compared. It is too time consuming to do this ‘by hand’.

## Ad 2 Additional features

- It would be helpful to have a *triangle comparison wizard*, according to the scheme shown in Figure 9. In this way one could easily compute Kappa statistics relative to a ‘Reality map’.
- For some analyses it would be interesting to aggregate both maps to a coarser grid. After that, we want to analyse differences by map statistics. However, in doing so, much information is lost by this aggregation: e.g. users have to decide for each coarser grid which category is dominant. Moreover, the user has to be careful that the frequency of categories for *the whole map* is not affected. The SPAT tool has been developed by RIKS just for this purpose. However, there are methods to aggregate categories proportionally to coarser grids without losing (much) information. It would be an interesting extension of the software to implement these techniques.

- If users are dealing with maps on an interval or ratio scale, and want to apply the fuzzy-map comparison approach, they have to transform the maps to ordinal maps first. There are possibilities to avoid this transformation and to apply the fuzzy approach *directly* to the interval/ratio maps. Such an approach would avoid the less elegant transformation step. The loss of information would also be avoided.
- It would make the MCK software all round if single map statistics could be evaluated (section 2.6). Then, we could perform all five map-comparison techniques described in Chapter 2. Especially interesting would be: (i) the addition of a fragmentation index and (ii) the enrichment factors, described in Appendix A.
- There are methods to compare maps with *unequal legends*. In some cases a re-classification of categories would be sufficient to get the legends equal. However, this will not be sufficient in all cases. A more advanced example is given in RIKS (2003c), where legends have been made equal by use of a contingency table. However, further experimentation is needed here. The feature would be helpful for a number of NEAA projects.



*The MAP COMPARISON KIT software is a product that deserves congratulations. It is a 'state of the art' map-comparison tool that will find its way to many users inside and outside RIVM. The software will be made publicly available on the RIVM website early 2004 ([www.rivm.nl](http://www.rivm.nl)). However, we still have comments and wishes for making the software even more complete than it is now, as we point out in this Chapter.*

*Photo: Advertisement campaign – Gemjobs.*

- More research should be done to Hierarchical fuzzy pattern matching. This method is promising. However, more should be known about the strong and weak sides of the approach.

### **Ad 3 Integration with other GIS software**

Using the MAP COMPARISON MODULE described in RIKS (2003b), the user can perform all MCK computations within her or his own GIS environment. Such an approach has the advantage that users do not switch from one software application to the other. Fuzzy-result maps can also be shown directly on the screen with all additional map information (e.g. provincial borders, rivers, large cities, etcetera). A simple prototype for ArcGis-8 is available, but a user-friendly user shell would be of great practical importance.

## References

- Anselin, 1988. Spatial econometrics. Methods and models. Boston, Kluwer.
- Bakkenes, M., Zwart, D. de, and Alkemade, J.R.M., 2002. MOVE nationaal model voor de vegetatie, versie 3.2. RIVM report 408657006.
- Beusen, A.H.W. and Visser, H., 2003. De KaartbeeldTransformator. Additionele softwaretool voor de FuzzyKaartbeeldVergelijker. IMP memorandum 001/2003.
- Borsboom – van Beurden, J.A.M., Regt, W. de, and Schotten, K., 2002. Land Use Scanner: the continuous cycle of application, evaluation and improvement in land use modeling. Paper 42th congress of the European Regional Science Association. Dortmund, Germany.
- Carletta, J., 1996. Assessing agreement on classification tasks: the kappa statistic. Computational linguistics, 22(2).
- CBS (1993). Bodemstatistiek 1989. Centraal Bureau voor de Statistiek. Voorburg, the Netherlands.
- CBS (1997). Bodemstatistiek 1993. Centraal Bureau voor de Statistiek. Voorburg, the Netherlands.
- CBS (2000). Bodemstatistiek 1996. Centraal Bureau voor de Statistiek. Voorburg, the Netherlands.
- Cheng, T., Molenaar, M., and Lin, H., 2001. Formalizing fuzzy objects from uncertain classification results. International Journal of Geographical Information Science, 15(1), 27-42.
- Cressie 1993. Statistics for spatial data. Wiley and Sons, New York.
- Dekkers, A.L.M., 2001. S-PLUS voor het RIVM. Krachtig statistisch software gereedschap. RIVM-rapport 422516001.

- De Nijs, A.C.M., Crommentuijn, L., Farjon, J.M.J, Leneman, H., Ligtoet, W., Niet, R. de, Schotten, K. and Schilderman, C.B.W., 2002. Vier scenario's van het landgebruik in 2030. Achtergrondrapport bij de Nationale Natuurverkenning 2, 2000 – 2030. RIVM rapport 408764003, Bilthoven, the Netherlands.
- Deutsch, C.V. and Journel, A.G., 1998. GSLIB. Geostatistical software library and user's guide. Oxford University Press.
- Engelen, G., R. White and T. de Nijs, 2003. The Environment Explorer: Spatial Support system for Integrated Assessment of Socio-economic and Environmental Policies in the Netherlands. Integrated Assessment, accepted.
- Fielding, A. H. and Bell, J. F., 1996. A review of methods for the assessment of prediction errors in conservation presence/ absence models. *Environmental Conservation*, 24(1), 38-49.
- Fisher, P., 2000. Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, 113(1), 7-18.
- G. M. Foody, 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80, 185-201.
- Goovaers, P., 1997. Geostatistics for natural resources evaluation. Oxford University Press.
- Hagen, A., 2002a. Comparison of maps containing nominal data. Report commissioned by National Institute for Public Health and the Environment (RIVM). Maastricht: Research Institute for Knowledge Systems.
- Hagen, A., 2002b. Multi-method assessment of map similarity. In: Proceedings of the 5th AGILE Conference on Geographic Information Science, 171-182. Palma (Spain), April 2002.
- Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps, *International Journal of Geographical Information Science*, 17(3), 235-249.
- Hilferink, M. and Rietveld, P., 2001. Een nadere uitwerking van het RuimteScanner model. In: Scholten *et al.*, 2001.

- Kaluzny, S.P., Vega, S., Cardoso, T.P. and Shelly, A.A., 1998. S+ Spatial Stats. User's manual for Window and Unix. MathSoft, Inc.
- Klein Goldewijk, K., 2001. Estimating global land use change over the past 300 years: the HYDE database. *Global Biochemical Cycles*, 15(2), 417-433.
- Klein Goldewijk, K. and Ramankutty, N., 2003. Land cover change over the last three centuries due to human activities: the availability of new global data sets (in print).
- Kuhnert, M., 2002. Map Comparison Survey. Retrieved from the Internet at <http://www.uvm.edu/giee/AV/Maps/> (June 26, 2002).
- Lantz, C. A. and Nebenzahl, E., 1996. Behavior and interpretation of the K statistic: Resolution of two paradoxes. *Journal of clinical epidemiology*, 49(4), 431-434.
- Maclure, M., & Willet, W. C., 1987. Misinterpretation and misuse of the kappa statistic. *American Journal of Epidemiology*, 126(2), 161-169.
- Maxwell, W. E., 1977. Coefficients of agreement between observers and their interpretation. *British Journal of Psychiatry*, 130, 79-83.
- Metternicht, G., 1999. Change detection assessment using fuzzy sets and remotely sensed data: an application of topographic map revision. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(4), 221-233.
- MNP, 2000. Nationale Milieuverkenning 5, 2000 - 2030. Rijksinstituut voor Volksgezondheid en Milieu. ISBN 90-140-7189-2. Samson bv. Alphen aan de Rijn, the Netherlands.
- MNP, 2001. Who's afraid of red green and blue? Toets van de Vijfde Nota Ruimtelijke Ordening op ecologische effecten. RIVM report 711931005, Bilthoven, the Netherlands.
- MNP, 2002a. Natuurverkenning 2. Rijksinstituut voor Volksgezondheid en Milieu. Kluwer. Alphen aan de Rijn, the Netherlands.
- MNP, 2002b. Natuurbalans 2002. Kluwer, Alphen aan de Rijn.

- MNP, 2002c. Ecologische effectberekening voor de 2e Nationale Natuurverkenning. Terrestische systemen. RIVM rapport 408664002, Bilthoven, the Netherlands.
- MNP, 2003. Milieubalans 2003. Het Nederlands milieu verklaard. Kluwer, Alphen aan de Rijn.
- Monserud, R. A and Leemans, R., 1992. Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, 62, 275-293.
- Myers, D. and A. Kitsuse, 2000. Constructing the future in planning: a survey of theories and tools. *Journal of Planning Education and Research* 19, 221-231.
- Pontius Jr., R. G., 2000. Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering & Remote Sensing*, 66(8), 1011-1016.
- Pontius Jr., R.G. and Schneider, L. C., 2001. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment*, 85(1-3), 239-248.
- Pontius Jr., R.G., 2002. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering & Remote Sensing* 68(10), 1041-1049.
- Pontius Jr., R.G., 2003. Methods to compare the accuracy of different types of land use change models. *Proceedings Framing Land use Dynamics. International Conference 16-18 April 2003, Utrecht University, the Netherlands.*
- Power, C., Simms, A. and White, R., 2001. Hierarchical fuzzy pattern matching for the regional comparison of land use maps. *International Journal of Geographical Information Science*, 15(1), 77-100.
- Ramankutty, N. and Foley, J.A., 1998. Estimating historical changes in global land cover: croplands from 1700 to 1992. *Global Biogeochemical Cycles* 13(4), 997-1027.

- Raziei, A. and Evers, W., 2001. Afstemming ruimtelijke gegevens voor de Leefomgevings-Verkenner en RuimteScanner. CIM-report M009/01. RIVM, Bilthoven, the Netherlands.
- RIKS, 2003a. Map Comparison Kit. User manual. RIKS report, July 2003.
- RIKS, 2003b. Map Comparison Module. Reference manual. RIKS report by I. Uljee, July 2003.
- RIKS, 2003c. LeefOmgevingsVerkenner. Expost analyse: 1970-2000. RIKS report by A. Hagen and M. van der Meulen, July 2003.
- Scholten, H.J., Velde, R.J. van de, and Borsboom - van Beurden, J.A.M., 2001. RuimteScanner: informatiesysteem voor de lange termijnverkenning van ruimtegebruik. Netherlands Geographical Studies 242. Edited by the Free University of Amsterdam.
- Schotten, K. Goetgeluk, R., Hilferink, M., Rietveld, P. and Scholten, H., 2001. Residential construction, land use and the environment. Simulations for the Netherlands using a GIS-based land use model. In: Environmental Modelling and Assessment, 6, 133-143.
- Verburg, P., Ritsema van Eck, J., De Nijs, T., Dijst, M., and Schot, P., 2003. Determinants of land allocation in the Netherlands. Environment and Planning B (accepted).
- Verburg, P., De Nijs, T., Ritsema van Eck, J., Visser, H. and De Jong, K., 2003. A method to analyse neighborhood characteristics of land use patterns. Computers, Environment and Urban Systems (in press).
- White, R., Engelen, G., Uljee, I., Lavalle, C. and Ehrlich, D., 2000, Developing an Urban Land use Simulator for European Cities. In: Proceedings of the 5th EC-GIS Workshop held in Stresa, Italy 30-31 June 1999, edited by E. Fullerton, (Ispra, Italy: European Commission, Joint Research Centre), pp. 179-190.
- White, R. and Engelen, G., 1993. Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environment and Planning A, 25, 1175-1199.
- Winter, S., 2000. Location similarity of regions. ISPRS Journal of Photogrammetry & Remote Sensing, 55(3), 189-200.

Zadeh, L., 1965. Fuzzy sets. *Information and Control*, 8, 338-353.

## Appendix A Neighbourhood characteristics

by P. Verburg and K. de Jong

We define a measure to characterize the over- or under representation of different land-use types in the neighbourhood of a location. This measure, the **enrichment factor** ( $F$ ), is defined by the occurrence of a land-use type in the neighbourhood of a location relative to the occurrence of this land-use type in the study area as a whole following:

$$F_{i,k,d} = \frac{n_{k,d,i} / n_{d,i}}{N_k / N} \quad (9)$$

$F_{i,k,d}$  characterizes the enrichment of neighbourhood  $d$  of location  $i$  with land-use type  $k$ . The shape of the neighbourhood and the distance of the neighbourhood from the central grid-cell  $i$  is identified by  $d$ .

**Figure 36** shows the shape of the neighbourhoods used in this study:  $n_{k,d,i}$  is the number of cells of land-use type  $k$  in the neighbourhood  $d$  of cell  $i$  and  $N_k$  the number of cells with land-use type  $k$  in the whole raster.  $N$  is the total number of cells.

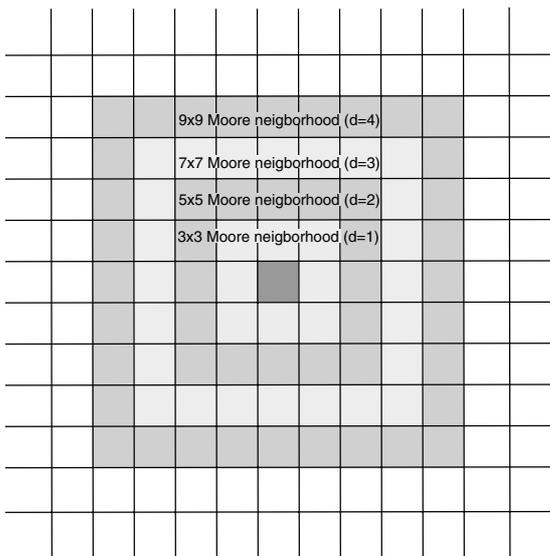


Figure 36 Configuration of neighbourhoods.

So, if the neighbourhood of a certain grid cell contains 50% grass, whereas the proportion grass in the country as a whole is 25%, we characterize the neighbourhood by an enrichment factor of 2 for grassland. When the proportion of a land-use type in the neighbourhood equals the national average, the neighbourhood is characterized by a factor of 1 for that land-use type. An underrepresentation of a certain land-use type in the neighbourhood results in an enrichment factor between 0 and 1.

This neighbourhood characteristic results for each grid cell  $i$  in a series of enrichment factors for the different land-use types ( $k$ ). The procedure is repeated for different neighbourhoods located at different distances ( $d$ ) from the grid cell so as to study the influence of distance on the interaction between land-use types. In this study we used square neighbourhoods (Moore neighbourhood, see Figure 36).

The average neighbourhood characteristic for a particular land-use type  $l$  ( $\bar{F}_{l,k,d}$ ) is calculated by taking the average of the enrichment factors for all grid cells belonging to a certain land-use type  $l$ , following:

$$\bar{F}_{l,k,d} = \frac{1}{N_l} \sum_{i \in L} F_{i,k,d} \quad (10)$$

where  $L$  is the set of all locations with land-use type  $l$  and  $N_l$  the total number of grid cells belonging to this set.

Confidence intervals for testing the significance of this characteristic cannot be calculated following standard procedures. Spatial autocorrelation causes all statistical tests to be biased and is therefore inappropriate for determining the significance of the calculated characteristics. For this reason we characterized the variability by the standard deviation and studied regional variability in the characteristics by comparing aggregated results with region-specific results. The variance is defined by:

$$s_{l,k,d}^2 = \frac{1}{(N_l - 1)} \sum_{i \in L} (F_{i,k,d} - \bar{F}_{l,k,d})^2 \quad (11)$$

### **Example**

In the following we will illustrate the calculation of neighbourhood characteristics. The examples given are for a 500 x 500 m version of the 1989 CBS land-use map. For details we refer to Verburg *et al.* (2003). For purposes of simplicity we have reduced the number of categories from 16 to 9, as summarised in **Table 16**.

Table 16 Land-use classification applied in this analysis

Land use type	Description
Other agriculture	Agricultural land not belonging to greenhouses, grassland or arable land, includes horticulture, orchards etc.
Grassland	Grasslands, incl. semi-natural grasslands
Arable land	All arable lands
Greenhouses	Greenhouses
Residential areas	Residential areas and social-cultural facilities, incl. houses, roads within residential areas, schools, hospitals, churches etc.
Industrial/ commercial	Mining areas, industries, harbours, shopping malls, prisons and all service industries
Forest/nature	Forests and natural areas, including peat areas, swamps, heather etc.
Recreation	Parks, sport fields, kitchen garden complexes, campsites etc.
Airports	Airports

Figures below show the enrichment factors for the categories: ‘recreation’ (**Figure 37A**) and ‘residential’ (**Figure 37B**). The graphs present the average enrichment factor ( $\bar{F}_{l,k,d}$ ) as a function of the distance ( $d$ ; see Figure 36). The enrichment factor is presented on a logarithmic scale to obtain an equal scale for land-use types occurring more than average in the neighbourhood (enrichment factor  $> 1$ ) and land-use types occurring less than average in the neighbourhood (enrichment factor  $< 1$ ).

For all land-use types considered, the neighbourhood tends to become less specific as the distance from the central cell increases. When values are close to 0, the composition of the neighbourhood is similar to the land-use composition of a random selection of grid-cells. At a distance of approximately 10 cells ( $d=10$ , with neighbourhood located between 5 and 7 km from the central cell), large deviations from the random composition no longer occur.

The most specific neighbourhood characteristics are found for neighbourhoods smaller than 5 units (between 2.5 and 3.5 km from the location). For most land-use types, the

neighbourhoods are most enriched by the land-use type itself over the whole range of neighbourhoods studied.

The positive enrichment of the **residential area** with industrial/commercial land use (blue line in Figure 37A) indicates that these land-use types are preferably allocated to locations close to residential areas, where a labour force and a market for products are available. However, residential areas and industrial compounds are preferably separated by some distance because of the potential consequential negative impact of industrial compounds on livelihood. This is represented by the somewhat lower enrichment of the neighbourhood located near an industrial area at a short distance from the residential area than the enrichment of such a neighbourhood at a larger distance from the residential area.

Another striking result, characteristic for the Dutch landscape pattern, is the negative interaction between a residential and *forest/nature* areas (yellow line) separated by short distances that turns into a slightly positive interaction at larger distances. Historically, arable land and grassland has surrounded settlements in the central and eastern part of the Netherlands, while, at a greater distance, wild lands were used for fuel wood and extensive grazing (heather fields). These patterns are still clearly visible in large parts of the Netherlands and explain the neighbourhood characteristics. This pattern may be reinforced by planning policies in the last decades aimed at protection of open areas, seen as characteristic for Dutch landscapes, This was done by re-directing residential development to more sheltered areas, where it is not visible from such a large distance.

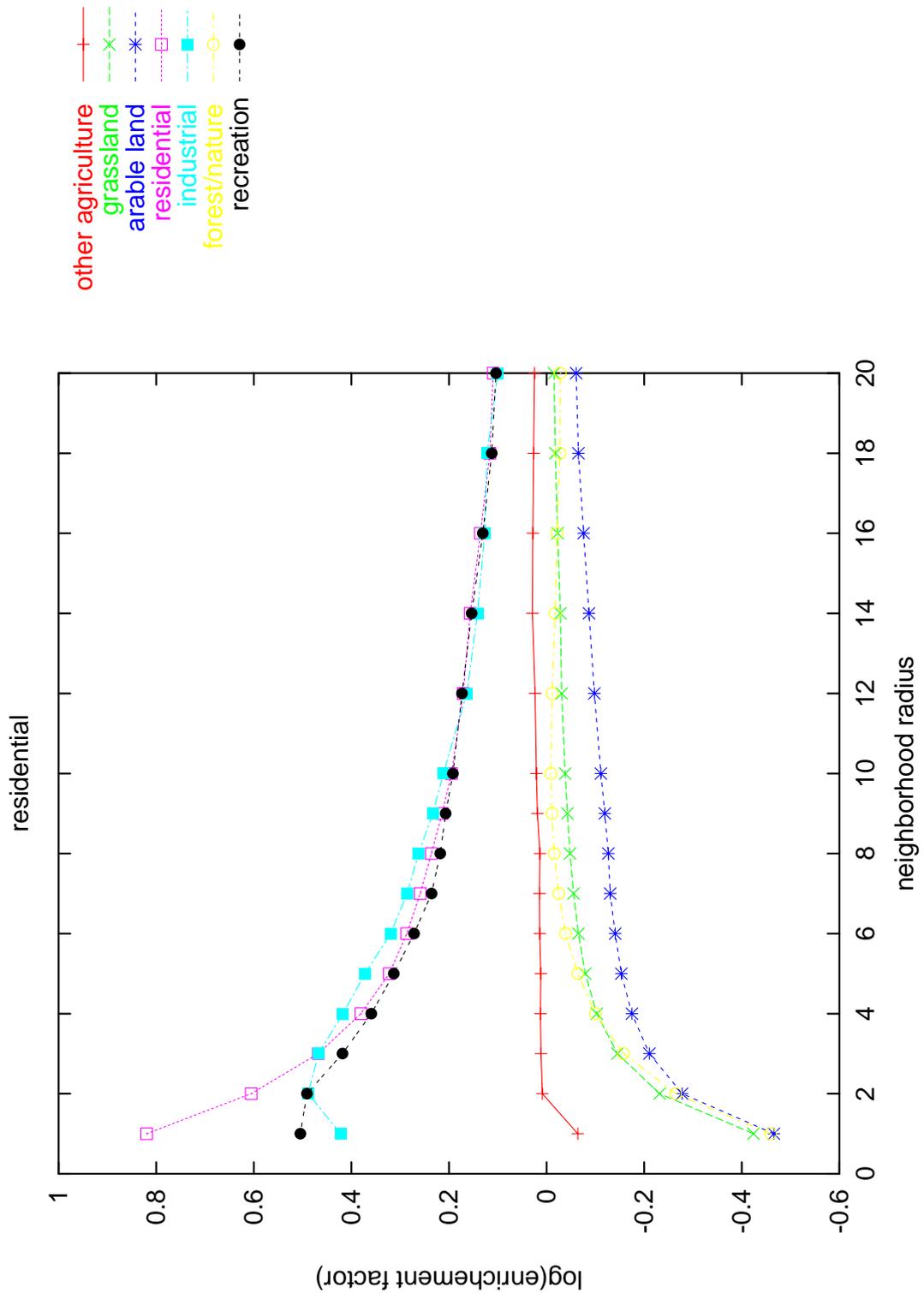
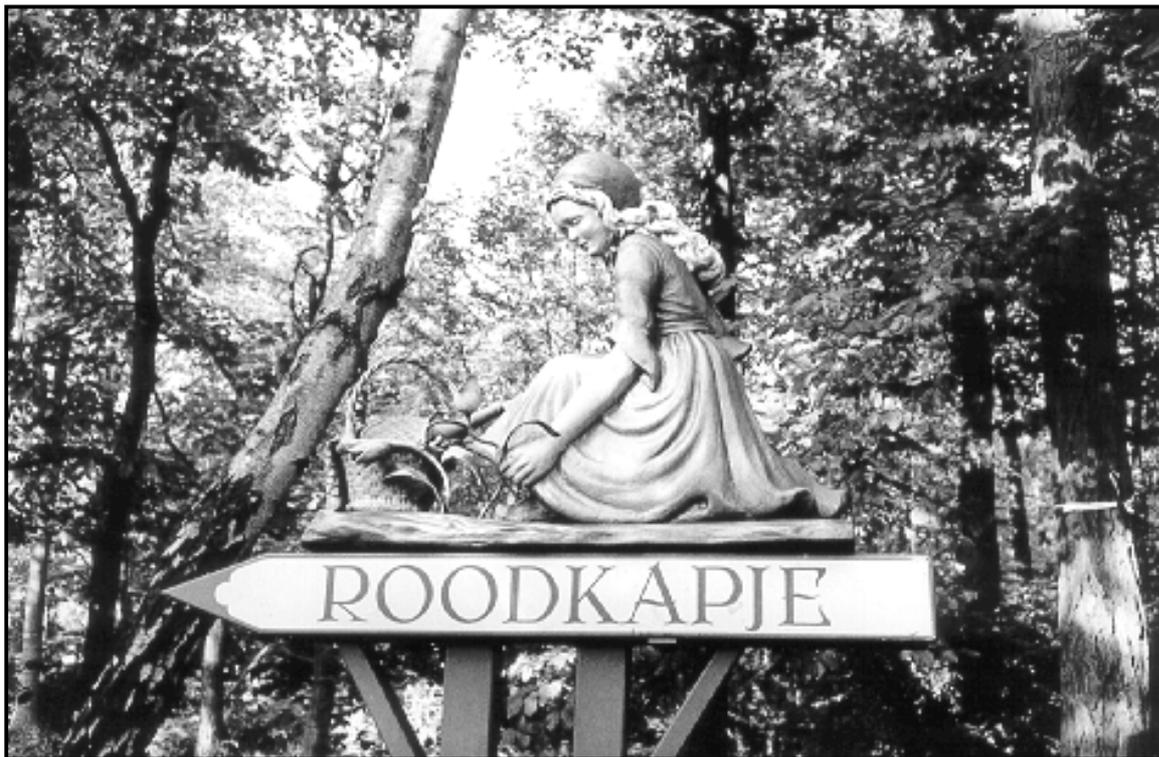


Figure 37A Logarithm of enrichment factors (y-axis) against neighborhood radii ( $d=20$  equals 10 km).

The enrichment factors are calculated for 'residential' versus seven other categories, including 'residential' itself.

**Recreation** (Figure 37B) has a high positive spatial autocorrelation for a small neighbourhood; however, industrial/commercial areas are, at some distance, more dominant in the neighbourhood, indicating the location of recreational and industrial/commercial areas at the edge of towns and cities.

For more details, and implications for cellular-automata model calibration, please refer to Verburg *et al.* (2003).



*Figure 37B shows the enrichment factors for 'recreation' and 'forest/nature' (yellow line). These enrichment factors show no specific attraction (positive values) or repulsion (negative values) for distances up to 10 km ( $d=20$ ). 'Recreation' versus 'industry' or 'residential' show strong attraction for distances of 1 to 5 km.. Clear repulsions are found for 'grassland' and 'arable land'.*

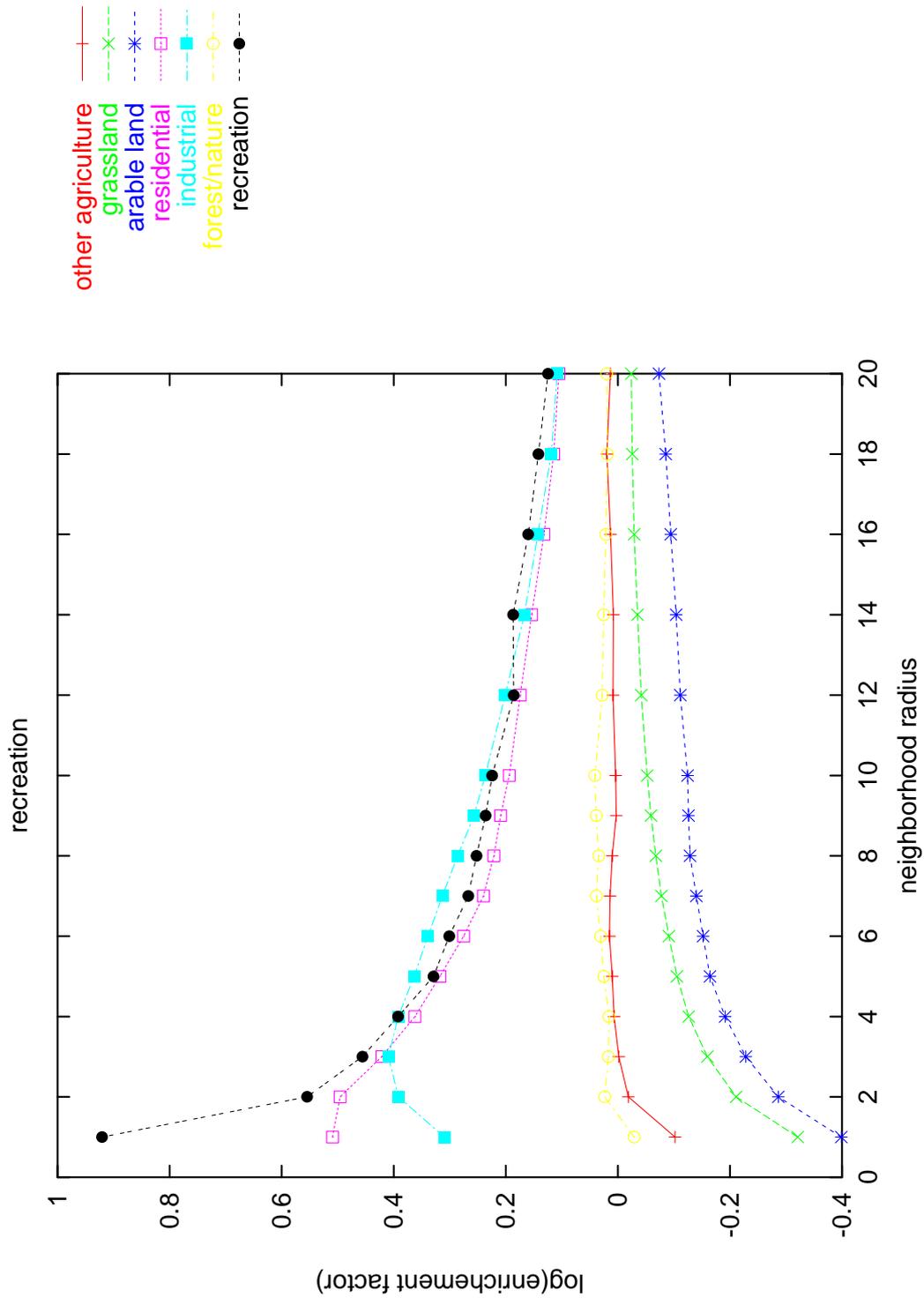


Figure 37B Similar to Figure 37A, but now for 'recreation' versus seven other categories, including 'recreation' itself.

## Appendix B Map transformations

### B.1 MAPTRANSFORMER

In practice not all maps exist in the right format for the MCK. For example, the categories in a set of nominal maps should be integer, starting at value '0', without any subsequent numeric gaps. These requirements are not always met in practice. Categories may start at '1', or a number of categories may be missing in one or more maps. Another method would be to combine a number of categories into one new category, and analyse map similarity afterwards. This operation cannot be performed within the MCK software.

Arthur Beusen of RIVM's Information Services and Methodology unit of the Environmental Assessment Agency (IMP/MNP in Dutch) developed a Fortran program, called the MAPTRANSFORMER for these types of transformations. The MAPTRANSFORMER runs on ArcInfo ASCII files (files with extension '.asc').

The software has the following options:

- testing for correctness of the input format. Nominal categories should start at '0' and end at 'N-1', with N the total number of categories in the set of coherent maps.
- renumbering of all maps if the test above so indicates.
- transformation of maps from interval or ratio scale to an ordinal scale. The classes may be filled automatically (two options) or chosen by hand.
- combining a number of categories into one new category.
- applying a mask file to each of the maps. This mask file contains ones at relevant grid points and zeros at grid points that should be left out of the comparison analysis.

A strong point of the software is that all operations are performed *simultaneously* to a set of raster maps (in general, classification within a GIS application is performed map by map, leading to wrong inferences). The user manual of the MAPTRANSFORMER is given in Beusen and Visser (2003).

It should be noted that the third and fifth options of the MAPTRANSFORMER may also be performed within the MCK. However, it may be advantageous to transform maps a priori. For example, if a mask file is used within the MCK, and if we want to copy a masked map to Word, the map is copied **without** the mask.

## B.2 Linking maps to S-PLUS for additional statistical analyses

Map comparison may be based on single map statistics, as we have seen in Chapter 2. In fact, one can perform a wide range of statistical analyses on single maps. And any statistic can be of potential use in map comparison. For example, one may want to know if there are spatial trends in one or more categories. In this case, it may provide insight to compare maps *after* removal of such trends. This procedure is very common in time-series analyses where signals are split into a trend (the low-frequency part of the signal) and residuals (the high-frequency part of the signal).

Or one might want to know if categories attract or distract each other as a function of distance. Such attraction/distract functions may help to validate cellular automata models, such as the Environment Explorer (Verburg *et al.*, 2003). A great number of map statistics are available within the statistical software package S-PLUS (Module S+ Spatial Stats in Kaluzny *et al.*, 1998).

Arnold Dekkers (IMP/MNP) programmed two S-PLUS functions for import and export of ArcInfo ASCII files (the files with extension '.asc'). The import function is 'f.arcinfo.in', the export function is 'f.arcinfo.out'. The function 'f.df2ai' is a help function within the export function. Furthermore, a global parameter, ArcInfoDIR, has to be filled with the directory path to the '.asc'-files, e.g.: ArcInfoDIR <- 'c:\\temp\\ArcInfo\\'.

The ASCII code containing the four S-PLUS objects described above, is 'splus.txt'. If this ASCII file is placed in c:\\temp\\ArcInfo, the object can be read into S-PLUS using the statement 'source('c:\\temp\\ArcInfo\\splus.txt')'. The S-PLUS line 'ls()' will show the contents of the working directory, including the four objects mentioned above.

A description of the functions and their respective argument lists is given in the S-PLUS help functions. Type 'help(f.arcinfo.in)' or 'help(f.arcinfo.out)'.

General information on S-PLUS at RIVM is given in Dekkers (2001).



## Appendix C Human judgement and fuzzy-set map comparison

by A. Hagen

### C.1 Introduction

Map comparison methods are used to express the similarity between maps, since traditional map comparison methods do not distinguish small differences from large differences. This leads to comparison results that are not in accordance with human judgement (compare Figure 2). Human judgement of maps appears to have more nuances than existing automated computer-based methods. This chapter will provide heuristic proof that the Fuzzy Kappa statistic resembles human judgement better than the traditional Kappa or the percentage of agreement.

Percentage-of-Agreement is calculated by dividing the number of cells that are identical in both maps by the total number of cells. A refinement of the Percentage-of-Agreement is the Kappa statistic (see section 2.3.1). This statistic corrects the Percentage-of-Agreement, (PA), for the percentage-of-agreement that can statistically be expected from a random rearrangement of all cells in both maps (PE). It is calculated according to Equation (4). The Fuzzy Kappa statistic is given in section 2.4.4 and the Kappa statistic in section 2.3.1.

A flaw of both the Percentage-of-Agreement and the Kappa statistic is that both are based on a cell-by-cell map comparison, in which the difference between cells is crisp; the cells are either identical or they are not. Not taken into account is that some categories are considered more similar to each other than others. Nor does either statistic have a certain tolerance for small spatial differences that a human observer would have.

Map comparison based on Fuzzy-set theory, as described in sections 2.3 and 2.4 in more detail, is a relatively new alternative approach. It does take tolerance for small spatial errors into account and also considers similarity between categories.

In the next section we will show via a heuristic approach based on an Internet questionnaire that the Fuzzy Kappa statistic resembles human judgement more than other statistics.

## C.2 Comparison of the comparisons

A test case will be considered here to validate the assumption that Fuzzy-set map comparison does approach human judgement closer than the crisp alternatives. The case was available from an Internet survey conducted by Matthias Kuhnert (2002). Participants in the survey were presented ten pairs of nominal maps and asked to express the similarity of each pair as a number between 0 and 1. A total of 127 participants filled out the form. The average similarity of each pair was calculated. See **Figure 38** for a summary of the map and survey results, and the results of the map comparison.

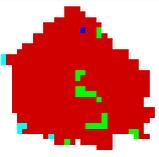
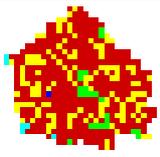
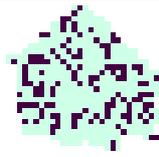
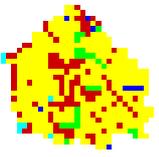
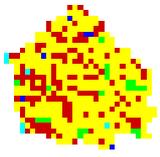
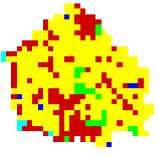
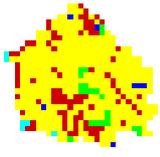
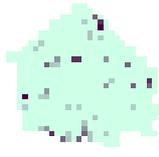
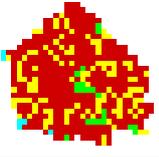
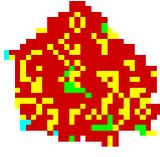
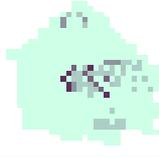
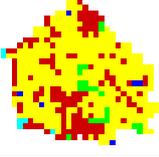
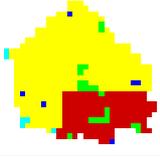
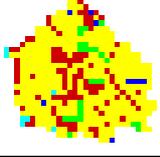
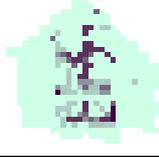
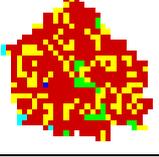
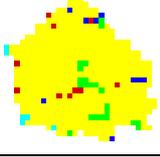
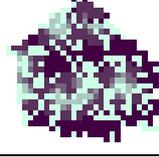
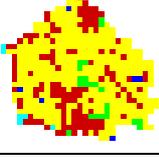
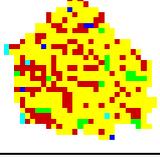
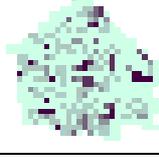
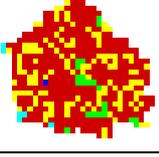
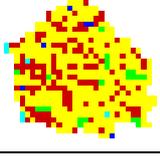
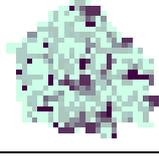
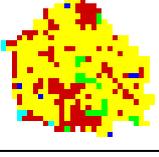
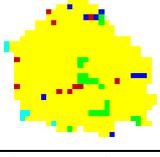
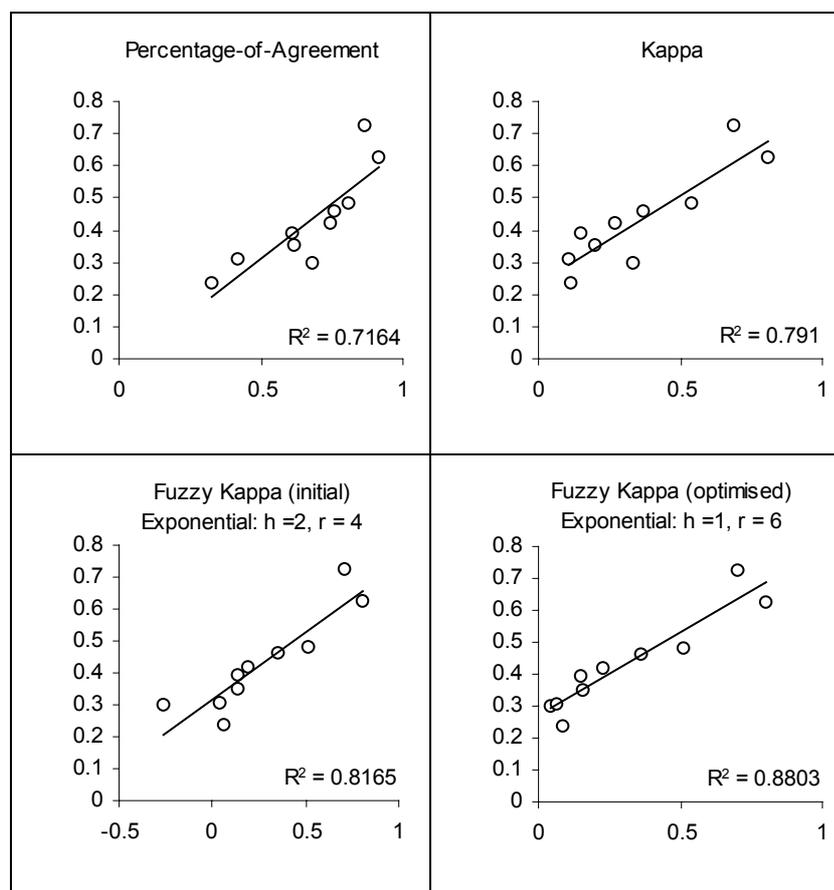
			Survey	0.42
			Percentage-of-agreement	0.75
			Kappa	0.27
			Fuzzy Kappa (optimized)	0.23
			Survey	0.39
			Percentage-of-agreement	0.61
			Kappa	0.15
			Fuzzy Kappa (optimized)	0.15
			Survey	0.62
			Percentage-of-agreement	0.91
			Kappa	0.81
			Fuzzy Kappa (optimized)	0.80
			Survey	0.73
			Percentage-of-agreement	0.87
			Kappa	0.69
			Fuzzy Kappa (optimized)	0.70
			Survey	0.30
			Percentage-of-agreement	0.68
			Kappa	0.33
			Fuzzy Kappa (optimized)	0.04
			Survey	0.48
			Percentage-of-agreement	0.81
			Kappa	0.54
			Fuzzy Kappa (optimized)	0.51
			Survey	0.24
			Percentage-of-agreement	0.33
			Kappa	0.11
			Fuzzy Kappa (optimized)	0.09
			Survey	0.35
			Percentage-of-agreement	0.62
			Kappa	0.20
			Fuzzy Kappa (optimised)	0.16
			Survey	0.31
			Percentage-of-agreement	0.42
			Kappa	0.11
			Fuzzy Kappa (optimised)	0.07
			Survey	0.46
			Percentage-of-agreement	0.76
			Kappa	0.37
			Fuzzy Kappa (optimised)	0.36

Figure 38 Ten pairs of maps found in the web survey and their comparison results.

The green-grey result maps indicate the level of similarity for each location. Darker cells show less similar levels.

The correlation between the similarity expressed by the participants and the comparison methods was evaluated (see **Figure 39**). The Fuzzy-set map comparison was performed for several fuzzy settings. The default fuzzy settings were first applied (exponential distance decay with a halving distance of two cells and a maximum radius of four cells), and a squared correlation of  $R^2 = 0.82$  was found for these settings. An  $R^2$  of 0.88 was found (for an exponential distance decay function with a halving distance of 1 cell and a maximum radius of 6 cells) after optimizing the settings. An  $R^2$  of 0.79 was found for the Kappa statistic and an  $R^2$  of 0.72 for the Percentage-of-Agreement.

We also calculated  $R^2$  for the survey results versus the K-location of Pontius, and the global similarity index of Power (not shown in Figures 12 and 13). We found an  $R^2$  of 0.51 for K-histo, and an  $R^2$  of 0.57 for the global similarity index.



*Figure 39 Scatterplots with the automated comparison results (x-axis) against the similarity expressed in the survey (y-axis). Also regression lines are given, along with corresponding  $R^2$  values.*

The cautious conclusion based on this fairly small test case is that the Fuzzy Kappa statistic approaches human judgement closer than the (crisp) Kappa and the Percentage-of-Agreement. The optimized version of Fuzzy Kappa shows the highest squared correlation ( $R^2 = 0.88$ ). This  $R^2$  is also higher than that of K-histo and the global similarity index ( $R^2 = 0.51$  and  $R^2 = 0.57$ , respectively).