GIMA

Geographical Information Management and Applications

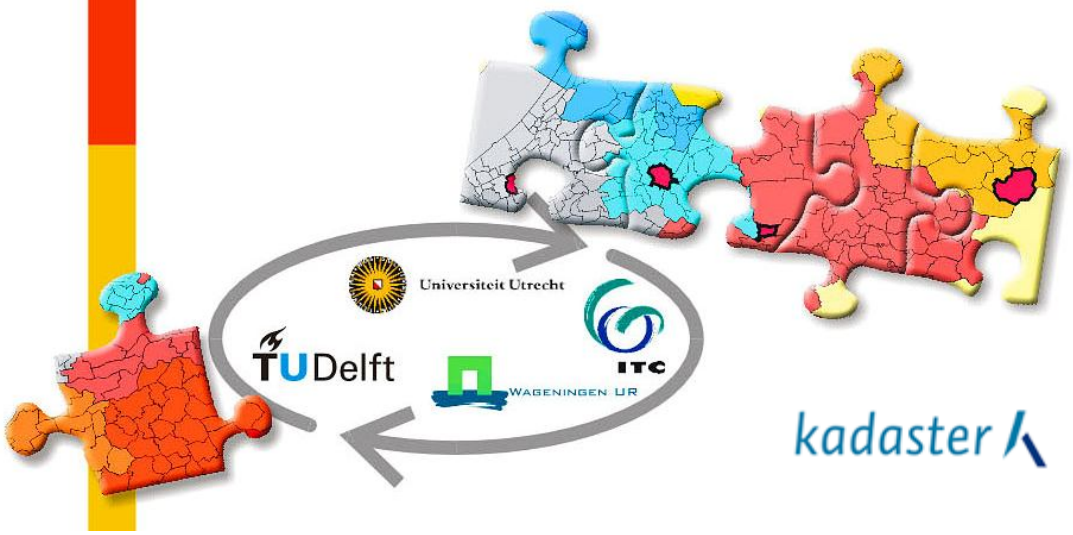
Exploring the possibilities of production of a culture codes dataset on cadastral parcel level by making use of external source datasets

MSc thesis

Supervisors:
Dr. Jantien Stoter (TU Delft)
Ir. Joop van Buren (Kadaster)
Drs. Gerard Leenders (Kadaster)

Professor:
Prof. Dr. Ir. Peter van Oosterom (TU Delft)

Author:
Emiel Dopper



Exploring the possibilities of production of a culture codes dataset on cadastral parcel level by making use of external source datasets

MSc thesis
November 2011

Author: Emiel Dopper

Supervisors: Dr. Jantien Stoter
Ir. Joop van Buren
Drs. Gerard Leenders

Professor: Prof. Dr. Ir. Peter van Oosterom

Preface

7 July 2010…..

This day can be called the birth date of the MSc thesis you are about to read. It was on this day that I replied to an email sent by Peter van Oosterom about an MSc thesis research project at the Dutch institution Kadaster.

The topic of the MSc thesis research project should be the update procedure of the Kadaster's culture codes dataset by using source datasets. An initial list of questions to be answered during the MSc thesis research project was included. Simple, short questions, but the answers would prove to be much more complicated and longer. As a result, the thesis did not have an easy childhood and it took a while for it to become a big boy.

As an adolescent, the thesis started to object against the expectations of its parent and others caring about him. The adolescence turned out to be a complicated period in the thesis's life, but was also the period in which important decisions had to be made concerning its future directions.

Finally, these decisions were made (not without hurdles) and after leaving adolescence behind, the thesis grew up quickly to become an adult. The result is the thesis lying in front of you.

I would like to thank my supervisors Jantien Stoter, Joop van Buren and Gerard Leenders for sharing with me their views and ideas on the thesis. Your comments have helped me to create this final product. Furthermore, I want to thank the professor Peter van Oosterom for giving useful comments on the research proposal and at the meetings.

I would also like to thank:

- Bart Maessen, Daniël te Winkel, Frank Kooij (all employees at Kadaster), Wim Hendriks and Janette Dijk (of the Ministry of Agriculture) for their input on the candidate source datasets;
- The Kadaster culture code project group for their input on the culture codes;
- Michael Karsters (Kadaster) for his support in obtaining the spatial datasets;
- Gerard Heuvink (Kadaster) for the FME Desktop introduction.

Special thanks go out to my parents, sisters, grandmother, other relatives, friends and (former and current) colleagues for their mental support and understanding. Without you I would not have been able to complete this thesis!

Finally, I would like to thank Teitur for his great show in Ekko (Utrecht) and his music. Your music has dragged me through the final weeks of my thesis project.

Emiel Dopper
Zeist, 7 November 2011

Abstract

Sharing and reusing spatial data becomes more and more important. One way of reusing shared spatial data is by creating a new dataset by making use of spatial data available in a combination of external datasets. This has been the central theme in this thesis, as its main objective has been to explore the possibilities of producing a culture codes dataset by making use of external source datasets. Furthermore, the quality of these source datasets, as well as their semantic, geometric and temporal characteristics have been taken into account.

The first step has been to describe and analyze a number of candidate source datasets, in order to be able to make a decision on which source datasets to use for the production of the culture codes. This included visual mappings between the classes and attributes of the candidate source datasets and the culture codes. Four source datasets have been selected (AKR, LKI, BAG and TOP10NL), with the choice for these datasets being mainly based on their contents, quality and accessibility.

After selecting these source datasets the visual mappings were translated into seventy-one decision rules for the production of the culture codes. Next, these decision rules were integrated into an extensive decision tree. The decision rules and the decision tree have been tested in a proof of concept. This proof of concept has been carried out in four case studies and quality analyses.

The results of this proof of concept have been quite satisfying, with 80,3% of all results being right. In order to further improve the results of the production method, a number of additional datasets has to be used and a few decision rules need to be investigated. Furthermore, the large amount of spatial data used for the calculations has caused serious performance problems. This should get special attention when deciding to apply an (automated) production of the culture codes by making use of the proposed production method. However, in general, the thesis shows that producing the culture codes dataset by making use of external datasets should be seen as a serious option.

Index

List of figures

List of tables

List of abbreviations

AKR	= Administratieve Kadastrale Registratie
BAG	= Basisregistratie Adressen en Gebouwen
BGT	= Basisregistratie Grootschalige Topografie
BRT	= Basisregistratie Topografie
BRP	= Basisregistratie Percelen
CBS	= Centraal Bureau voor de Statistiek
GBKN	= Grootschalige Basiskaart Nederland
ISO	= International Organization for Standardization
LADM	= Land Administration Domain Model
LGN	= Landelijk Grondgebruik Nederland
LKI	= Landmeetkundig en Kartografisch Informatiesysteem
LV	= Landelijke Voorzieningen
NEN	= Nederlands Normalisatie-Instituut
OBL	= obligatory
OPT	= optional
OWL	= Web Ontology Language
SDI	= Spatial Data Infrastructure
TC	= Technical Committee
UML	= Unified Modeling Language
WMS	= Web Map Service
WOZ	= Wet Waardering Onroerende Zaken

Chapter 1 – Introduction

The sharing and re-use of spatial datasets is becoming more and more important. An important reason for this development is that the creation of these datasets is, in most cases, a time-consuming and an expensive job (van Loenen et al. 2008). By making use of standards in data format and various technical solutions for sharing the data, it becomes easier to obtain and re-use spatial datasets created by others (Crompvoets 2009; de Man 2007; Feeney and Williamson 2000; Rajabifard and Williamson 2001). Furthermore, it is possible for users of these shared datasets to create their own data from these input datasets. This can be done in many different ways: by combining the spatial data with other data types (administrative, demographic, etc.), by extracting a subset from the source dataset, or by combining information from different shared spatial datasets in order to create a new dataset, etc.

The research presented in this document is an example of the third way of creating new data from shared spatial datasets: combining different spatial datasets in order to create a new dataset. Put shortly, the purpose is to use a number of available datasets to create a land use dataset for cadastral parcels, called the culture code dataset. This new culture code dataset should replace the existing dataset, which is produced manually by employees of the Dutch institution Kadaster.

This chapter serves as an introduction to the research project. It starts with an introduction into the tasks of the institution Kadaster (1.1) and the position of the culture codes in relation to these tasks (1.2). Then, it is explained why the Kadaster is interested in a new way of producing the culture code dataset – the motivation for the research project – (1.3). Section 1.4 provides information on the scope of the research project and presents the research questions to be answered after conducting the research. The methodology applied is presented in section 1.5. Finally, section 1.6 serves as a further reading guide by providing a short overview of the contents of the other chapters of this document.

1.1 – Land registration and cadastre

Palmer and McLaughlin (1996) explain that in democratic market-oriented societies there is a need for the registration of property rights. Such a registration is called a *land registration* and includes, according to the authors, among others rights related to ownership, charges and liens. McLaughlin and Nichols (1989, pp. 81-82) define land registration as “the process of recording legally recognized interests (ownership and/or use) in land”. Zevenbergen (2004) provides a similar definition by defining land registration as “a process of official recording of rights in land through deeds or title (on properties)” (p. 11). In general, one could state that land registration answers the questions “who?” and “how?” in figure 1.1 by registering owners and rights (titles) (Zevenbergen 2002, p. 29; 2004).

Strongly related to the concept of land registration is the concept of *cadastre*. McLaughlin and Nichols (1989, pp. 81-82) describe cadastre as “an official record of information about land parcels, including details of their bounds, tenure, use and value”. Zevenbergen (2004, p. 11) gives a similar definition of cadastre: “a methodically arranged public inventory of data concerning properties within a certain country or district, based on a survey of their boundaries”. This relates to the questions “where?” and “how much?” in figure 1.1 (Zevenbergen 2002, p. 29; 2004). As the location of the boundaries and the area of the property are stored in the cadastre, it is a geographical record, whereas the land registration solely contains administrative information. The relation between land registration and cadastre, and their position within a larger economical context, are visualized in figure 1.2.

This figure shows that cadastre, containing the location, size and boundaries of a parcel, is the smallest element in the economy at large, followed by land registration. They serve to explain and describe such economical actions as land transfer and land transaction, which are taking place within the broader contexts of a land market or the economy at large (Zevenbergen 2002, pp. 13-15).

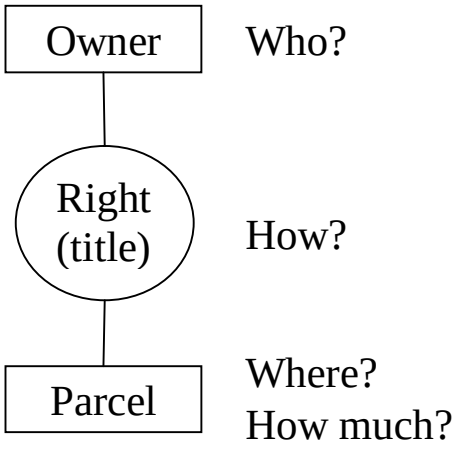


Figure 1.1: Core entities of land registration and cadastre. Source: Zevenbergen (2004, p. 12).

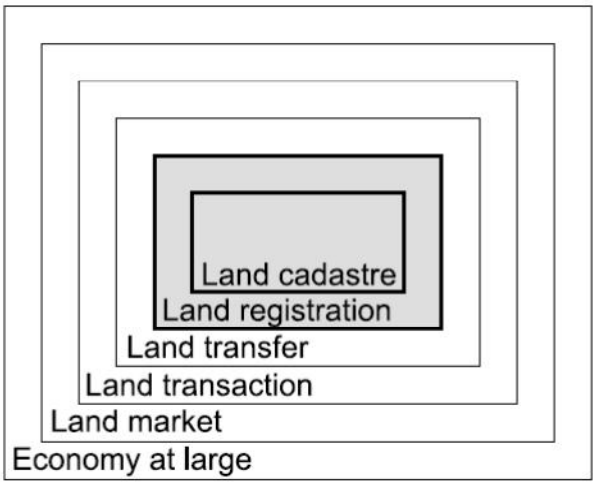


Figure 1.2: Land registration as a nested subset of the economy (Zevenbergen 2002, p. 14).

Zevenbergen (2002, 2004) argues that figure 1.1 is a static model of the system of land registration, which can only be used to store and find connections between owner, right and parcel. The author proposes to use a dynamic model of land registration, rather than the static model. Such a dynamic model is shown in the right part of figure 1.3. The static model can still be distinguished, as it is represented by the mushroom figure. However, the model has been extended to include two updating circles which are continuously going on:

- Land transfer: this can be explained as a change in the rights holder. For example: person A buys a parcel from person B, which makes person A the new rights holder for this particular parcel.
- Mutation: a change in the size of the cadastral parcel. In figure 1.3 subdivision represents the concept of mutation. However, subdivision (which is the splitting up of a parcel into a number of smaller parcels) is just one example of a cadastral mutation. Another example of a cadastral mutation is the opposite of subdivision:

'amalgamation'. In this case a number of parcels is combined to form one bigger new parcel.
(Zevenbergen 2002, pp. 106-109; 2004)

Finally, a few words on the usage of the terms cadastre and Kadaster in this document. In most countries land registration and cadastre are organized in separately functioning institutions. In the Netherlands, the two components are organized together within the institution Kadaster. Therefore, in this document, when referring to Kadaster, this includes both land registration and cadastre.

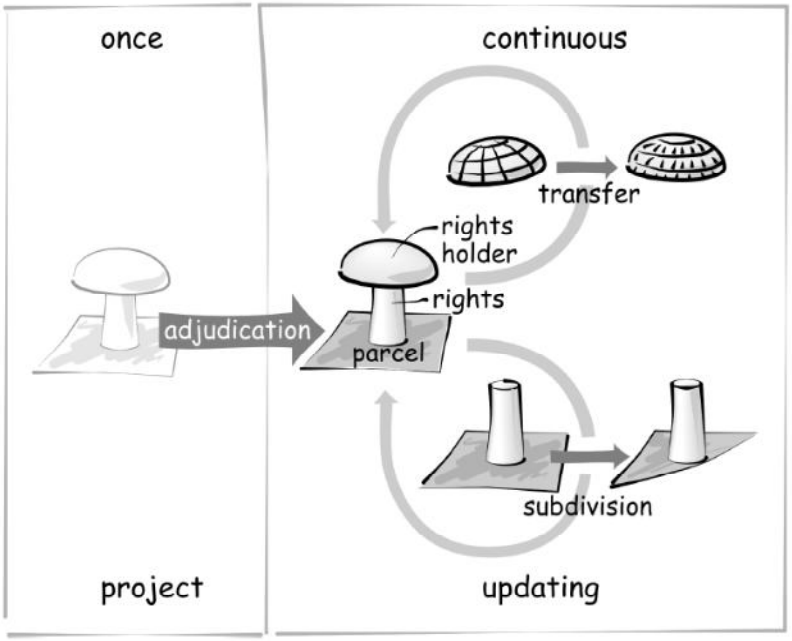


Figure 1.3: Dynamic model of the system of land registration (Zevenbergen 2004, p. 16).

1.2 – Culture codes

By adding the question "what?" to figure 1.1, one could explain what is located on a certain parcel (see figure 1.4, in red). In other words: this would describe the land use of a single parcel. This would then be the position of the culture codes as described in this document[1]. Land use is also included in the draft version of the LADM (Land Administration Domain Model) of the ISO. In the LADM land use is seen as an external class and five different types of land use – agriculture, housing, industry, nature and recreation – are specified (ISO TC 211/SC 2010, pp. 89-91). Figure 1.5 presents the UML class diagram[2] of the LADM, including the land use classes. The culture code dataset distinguishes between a much larger amount of land use classes (see appendix A for the full list of former and current culture codes).

[1] It should be mentioned here that the production of the culture code dataset, unlike land registration and cadastre, is not an official task of the Kadaster. However, as many internal and external products of the Kadaster are based on or make use of the culture code dataset, it plays an important role in the organization.
[2] UML (Unified Modeling Language) class diagrams give "an overview of a system by showing its classes and the relationships among them" (Miller 2003).

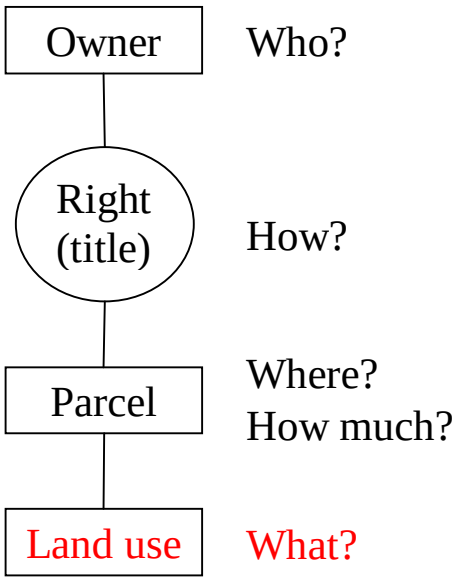


Figure 1.4: Core entities of land registration and cadastre, and the position of culture codes (taken from Zevenbergen (2004, p. 12) and edited by the author).

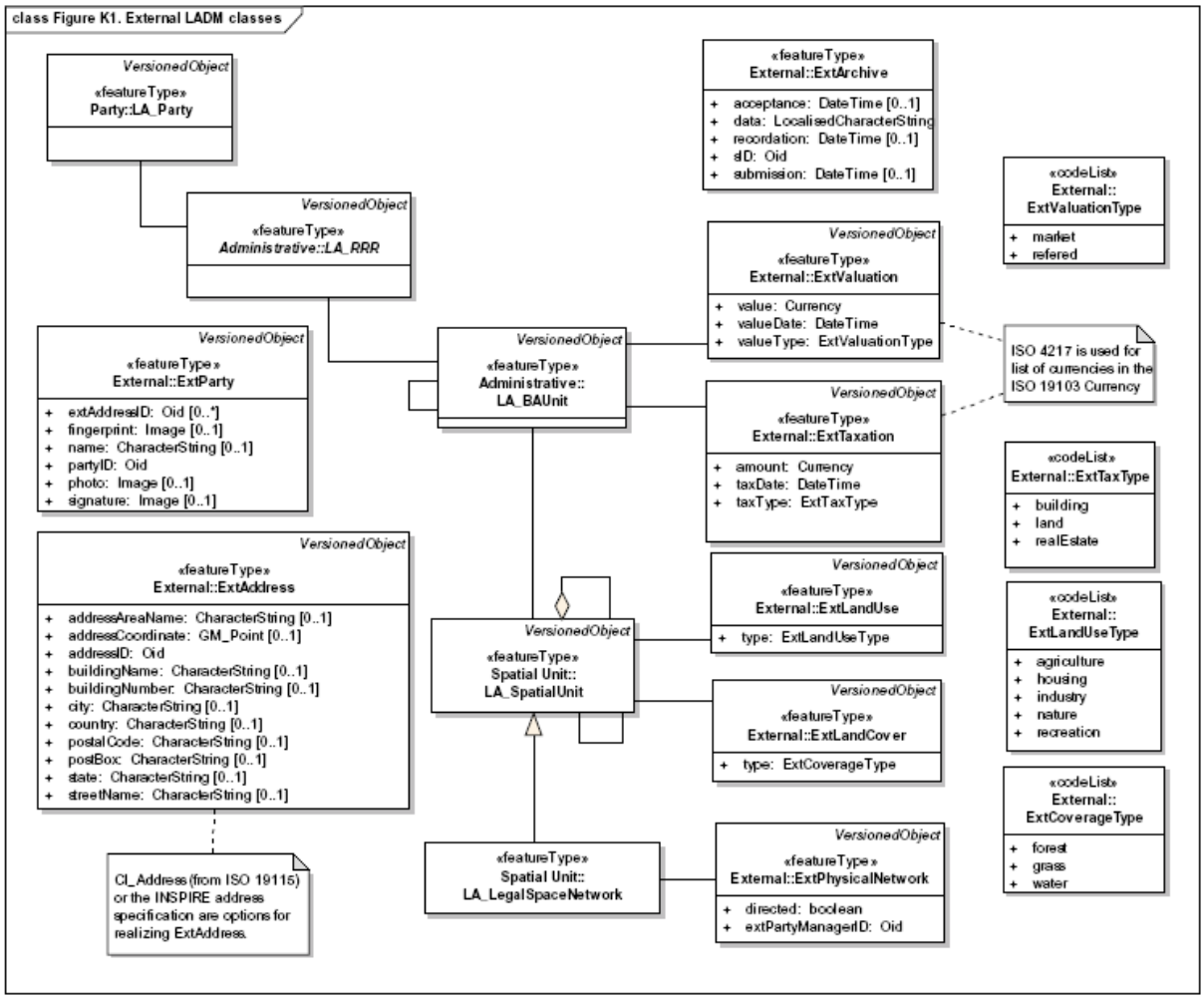


Figure 1.5: UML class diagram of the LADM (ISO TC 211/SC 2010, p. 89).

As mentioned above, the culture code is a code indicating the land use of a single parcel. Ellis (2010) argues that different scientific disciplines have different definitions of land use: "natural scientists define *land use* in terms of syndromes of human activities such as agriculture, forestry and building construction that alter land surface processes including biogeochemistry, hydrology and biodiversity. Social scientists and land managers define *land use* more broadly to include the social and economic purposes and contexts for and within which lands are managed (or left unmanaged), such as subsistence versus commercial agriculture, rented vs. owned, or private vs. public land" (Ellis 2010). In the context of the culture codes of this study, the second definition – of the social scientists and land managers – is most appropriate. Although "land use codes" or "parcel land use codes" might be clearer terms describing the contents of the culture codes than the current term (culture codes), it is chosen to stick to the term "culture codes" in this research project. The main reason for this is that the term is largely accepted within the institution Kadaster.

The culture codes can be seen as an example of a land use classification. *Land use classification* is defined by Mayhew (2004) as: "the analysis of land according to its use: agricultural, industrial, recreational and residential.[…] On occasion, the land may have more than one use, as in upland areas used for sheep farming and for recreation" (Mayhew 2004, p. 292). However, the culture codes have a predefined shape and size, as they represent the land use of cadastral parcels and the addresses located on these parcels. In the current situation, the culture codes are stored in the AKR, which is the administrative cadastral register that also contains the real estate rights established on the parcel. In AKR, it is possible to insert and change the land use information for cadastral parcels. To do so, one first has to select a specific parcel[3]. Second, the option for changing the '*cultuurgegevens*' (culture/land use indication) has to be chosen. Third, it is possible to select a culture indication best describing the built-up elements situated on the parcel (see figure 1.6). Two other options are available: (1) "special objects", in case the land use is not available in the list of culture indications, and (2) "only non-built elements", in case there are no built-up elements situated on the parcel. Fourth, one can choose a culture indication describing the non-built elements on the parcel, if there are any, or the option special objects (see figure 1.7). Fifth, when the option special objects has been chosen in one of the previous screens, in a new screen it is possible to insert a description of the land use on the parcel (free text). Finally, the addresses situated on the parcel are shown. In this screen one can add or remove addresses, if necessary. The culture indications are stored with each of the addresses located on the parcel, as well as with the parcels. When one has indicated there are built-up elements located on the parcel, it is obligatory to have at least one address in the addresses screen (Kadaster 2004a, 2004b).

After these steps, the AKR automatically produces the culture codes, based on the culture indications chosen during the procedure described above. Furthermore, the AKR defines, based on the choices during the procedure, a "*bebouwingscode*". There are four possible codes for the *bebouwingscode*:

1. Non-built-up;
2. Non-built-up and built-up;
3. Built-up;
4. Special properties. (Kadaster 2004b)

The *bebouwingscode* defines the number of culture codes added to a cadastral parcel. If the *bebouwingscode* is either "non-built-up" or "built-up", the cadastral parcel can have only one culture code. Cadastral parcels having the *bebouwingscode* "non-built-up and built-up" can possess two culture codes: one for the non-built-up objects and one for the built-up objects

[3] Every parcel has a unique code, made up of three sub-codes for: (1) municipality, (2) section within the municipality, and (3) the individual parcel.

located on them. In the current situation, cadastral parcels with special properties can have only one culture code, “Special properties”, which does itself not tell whether the cadastral parcel contains built-up or non-built-up objects. If the bebouwingscode “special properties” is chosen, the Kadaster employee can add free text to the culture code. As a result, this code is often chosen when a cadastral parcel contains more than one land use type or when the employee is not sure which land use type to choose.

In this thesis it is suggested to use “Special properties” not as a *bebouwingscode*, but as one of the possible culture codes. One reason for this is that special properties can refer both to built-up and non-built-up objects. It is also possible that cadastral parcels have a built-up culture code “Recreation – Sport” and special properties as non-built-up objects (something not listed in the culture codes list). Another reason for treating “Special properties” as a culture code is that it is also present in the list of culture codes of Kadaster (see appendix A). This suggests that “Special properties” is not a *bebouwingscode*, but rather a culture code.

Figure 1.6: In screen MD 34 one can choose one of the listed culture indications for built-up elements or the options special objects or only non-built elements in the lower right corner (Kadaster 2004b).

Figure 1.7: In screen MD 35 one can choose one of the listed culture indications for non-built elements or the option special objects (Kadaster 2004b).

1.3 – Why a new way of producing the culture codes?

Until a couple of years ago the list of culture codes contained about eighty codes (see Appendix A), which has been brought down to a remaining number of about forty. These codes include, among others, “residential”, “terrain (nature)” and “activity (industry)”. A range of problems related to the culture codes forces the Kadaster to search for an improved way of producing and/or updating the codes. The most important problems are described in this section.

The production of the culture codes is a manual process. When a parcel is sold, the transaction document contains a description of the “land use” on the parcel, which is converted to one or more codes which are the culture codes. In the case the description in the current transaction document differs from the one in the previous document, at Kadaster the parcel is provided a new culture code (according to the procedure described in the previous section). One could call this an “ad hoc” and “occasional” way of production. When the land use of the parcel changes during the period in between two transactions, this is (usually) not observed by Kadaster and, therefore, the culture code is not changed. An example to illustrate the problem: a parcel with the culture code “land (new construction residential)” is sold in 1990. Two years later the construction of the house is completed. Only in 2010, when the house is sold, a new transaction document is written and the culture code is changed to “residential”. In the 18-year period between completion and selling of the house, the culture code has been ‘wrong’ (construction site instead of housing).

Furthermore, as the production of the culture codes is a manual process, it is subject to personal interpretation of the culture description in the transaction document. For example, a culture description “nature area with a number of small fens” could be interpreted by one employee as nature area and by another as fens. Again another employee could interpret it as water. This subjectivity strongly threatens the trustworthiness of the culture codes.

Next to the problems related to the up-dating process and the subjectivity involved in the production of the culture codes, there are problems concerning the continuity of the production. Kadaster cannot force notaries to include a culture code in the deeds or transaction files – as the production of the codes is not an official task of the organization. Kadaster is planning to introduce a standard transaction file in order to facilitate automatic data collection and one component of this file could be the choice of a culture code to describe the land use on the parcel. However, the notaries are not willing to limit their description of the sold/bought property to the codes, as the codes do in many cases not give a detailed and/or complete enough description of the actual situation. For example: the culture code “residential” does not tell the user of the codes the same as the culture description “large bungalow with swimming pool and garden”, which are essential notions for the notary.

A further problem is related to the procedure of changing/adding a culture description to a certain parcel. As described in the previous section, it is possible for employees of Kadaster to choose the option “special objects”, which allows them to enter free text as the culture description. This option is sometimes used to be able to enter either a more precise description or to enter more than one culture description, although it should only be used for those special objects not covered by the list of culture codes. As a result, when creating products by using the culture codes, a number of parcels will not be covered, which harms the quality – in terms of logical consistency, explained in section 2.3 – of the products.

Finally, and related to the problems described above, the general quality of the culture codes data set is unsatisfying. This low quality is, among others, caused by outdated data and subjective choices of the employee adding the culture code. As a large range of products are based on the data set of culture codes, this makes the quality of these products questionable: “rubbish in, rubbish out”.

Summing up, the following problems force the Kadaster to search for a new method of producing and/or updating the culture codes data set:

1. Ad hoc updating based on transaction files;
2. Subjectivity because of the personal interpretation of culture text in transaction file;
3. Continuity issues;
4. Inappropriate usage of the option “special objects”;
5. General low quality of the data set.

1.4 – Research scope and questions

This section gives a further introduction into the research project presented in this document. It describes the research scope and presents the research questions which are to be answered at the end of this document.

1.4.1Research scope

The entire project concerning the production of the culture codes dataset, by making use of external datasets, consists of four components. These four components are the following:

1. Research on the needs of users of the culture codes. This focuses on the question which culture codes are needed to fulfill the requests of clients of the Kadaster;
2. Research on the content and quality of a number of potential source datasets for producing the culture codes dataset. This includes conceptual mappings between the source datasets and the culture code dataset;
3. Research on the possibilities of a matching procedure between the source datasets and the cadastral parcels in order to be able to derive the culture codes. This includes questions concerning the semantic, geometric and temporal possibilities, as well as potential problems coming with this procedure;
4. Creating an automated method for the procedure developed in step 3.

A special internal project group of Kadaster is carrying out the first component of this list, as they have first-hand information on the requests coming in from clients. Furthermore, they can get in contact with clients when they need additional information on their needs. The research project presented in this document focuses on the second and third component from the above listed components. As a result, the research project does not include the creation of an automated method for the procedure developed and tested in this thesis. In order to be able to thoroughly investigate the content and quality of a number of potential source datasets, as well as the possibilities of using these datasets for a matching procedure with the cadastral parcels in order to produce the culture codes, and at the same time taking into account the time constraints of the research project, the fourth component is a topic for further research. The thesis does, however, include conceptual mappings which are translated into decision rules and a decision tree, which can be seen as a starting point for an automated production procedure for the culture codes dataset. Furthermore, the final chapter of this MSc thesis does include, next to the conclusions, a list of recommendations for further research and the next steps to be taken to produce the culture code dataset in an automated manner.
In short, therefore, the scope of this research project is two-fold:

- Research on the content and quality of a number of potential source datasets;
- Research on the semantic, geometric and temporal aspects of a matching procedure between these source datasets and the cadastral parcels in order to produce a culture code dataset.

1.4.2 Research questions

The main research question of this thesis is:

> *To what extent is it possible to automatically produce a culture code (land use) dataset by making use of external source datasets, when taking into account their quality as well as their semantic, geometric and temporal characteristics?*

Next to the main research question a number of sub-questions have been formulated. The answers to these sub-questions should, together, help to formulate an answer to the main research question. The sub-questions are as follows:

1. Which source datasets could be used for an automated production of the culture code dataset?
    a. What is the content of these source datasets?
    b. What are the temporal characteristics of these source datasets? Do the datasets contain historical attribute values?
    c. What is the quality of these source datasets?
    d. What is the accessibility of these source datasets?

2. To what extent is it semantically possible to use these source datasets for an automated production of the culture code dataset?
    a. Which conceptual mappings can be created between the source datasets and the culture codes?
    b. Which decision rules can be formulated for the production of the culture codes?
    c. To what extent do the attributes of the source datasets cover the culture codes?

3. To what extent is it geometrically possible to use these source datasets for an automated production of the culture code dataset?
    a. What are the geometric characteristics of the source datasets?
    b. What happens when performing an overlay procedure with the candidate source datasets and the cadastral parcel map?
    c. Which geometric hurdles/problems should be overcome for an automated production of the culture codes dataset to be feasible?

4. What are the follow-up steps to be taken in order to be able to (automatically) produce the culture codes dataset?
    a. Which additional geographical information is needed in order to be able to create all contemporary culture codes?
    b. Which next steps should be taken in order to introduce an automatic way of producing the culture codes dataset?
    c. Which issues require special attention when introducing such an automatic way of producing the culture codes dataset?

1.5 – Methodology

This section gives information on the research method applied to obtain the answers to the main research question and sub-questions formulated in the previous section (1.4). It does so by presenting the main research steps carried out and research techniques used during the

research project (section 1.5.1), and by listing the datasets and software packages (1.5.2) which are used.

1.5.1 Research steps and techniques

In order to be able to answer the main research question and the sub-questions, a number of research steps has to be carried out. Figure 1.8 shows a flowchart visualizing the different steps to be taken. This flowchart is followed by a short explanation for each research step and the research techniques applied.

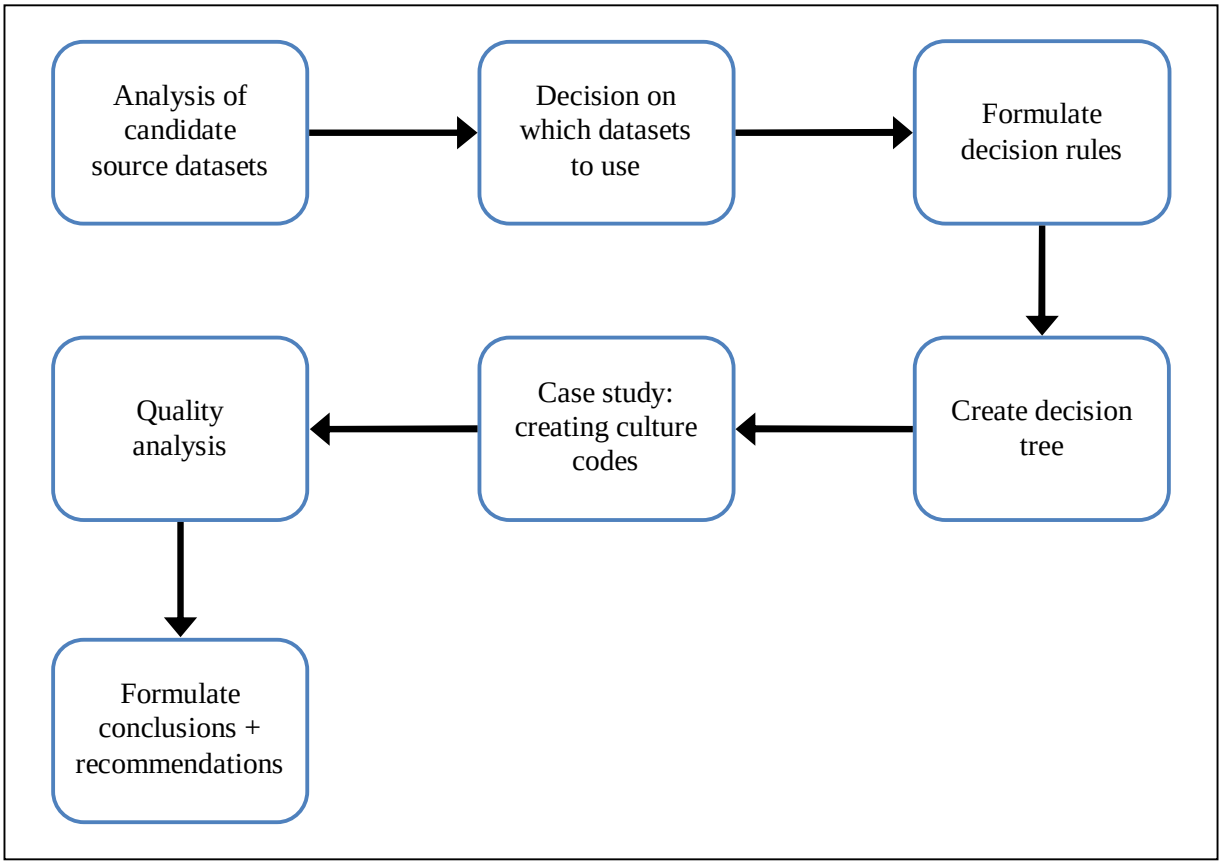


Figure 1.8: Flowchart for the research project.

As has been shown in the flowchart, seven main steps can be distinguished. These steps are:

- Analysis of candidate source datasets. Six candidate source datasets have been chosen (BAG, BRP, GBKN/BGT, TOP10NL, LGN and AKR/LKI). These dataset have been selected for different reasons. First, national spatial datasets are needed for the production of a country-wide culture code dataset. Second, the datasets should contain spatial information needed for the production of the culture code. The BAG dataset has, therefore, been chosen as it contains much useful information on buildings and the use of these buildings. The BRP and LGN have been selected as they contain detailed information on agricultural and nature land use types. The AKR/LKI dataset has been chosen for its information on cadastral parcel boundaries and apartment rights. Finally, the TOP10NL and GBKN/BGT datasets are selected because of their great variety of spatial information (especially the TOP10NL dataset). Third, these candidate source datasets are expected to have a quite good basic level of quality. For some of these datasets this quality should be guaranteed as they are part of the Dutch system of key registers.

The candidate source datasets are analyzed by making use of documentation on these datasets and by interviewing specialists on the datasets. The interviews are, however, seen as another source for the description and analysis of the datasets, next to the documentation. They are semi-structured, as a list of topics has been defined, but no standard questions are used. The interview topics are also used in the dataset analysis in this report. The following topics are distinguished:
    - General content of the dataset.
    - Geometry used in the dataset.
    - The production method used.
    - The up-to-dateness of the dataset.
    - Quality aspects/Trustworthiness of the dataset.
    - Listing the useful (attribute) classes of the source dataset and a visual mapping between these (attribute) classes and the culture codes.
    - The advantages/disadvantages of the dataset.
- Decision on which datasets to use. Based on the analysis of the candidate source datasets, a decision is made on which datasets should be used for the production of the culture codes in the remaining part of the research project.
- Formulate decision rules. Making use of the visual mappings between the useful (attribute) classes from the source datasets and the culture codes, a list of decision rules is formulated.
- Create a decision tree. The decision rules are used to create a decision tree for the production of the culture codes.
- Case study: creating culture codes. After formulating the decision rules and creating a decision tree, the proposed method is tested. This is done by calculating the culture codes for cadastral parcels, making use of shapefiles of (parts of) the source datasets. In ArcMap a number of analyses is carried out for these calculations. The different analysis steps are described in chapter 5.
- Quality analysis. The results of the testing procedure have been validated. This is done by making a comparison between a number of culture codes produced during the testing procedure and the real-world situation. The researcher has visited the four different testing areas. In every area, a number of parcels has been investigated, based on the culture codes assigned to them. Which parcels have been investigated, and the reasoning behind the choice of these parcels, is described in chapter 5.
- Formulate conclusions and recommendations. Finally, based on the analysis of the candidate source datasets, the decision rules and decision tree, and the test results, a number of conclusions and recommendations are formulated. The conclusions give an answer to the research questions presented in section 1.4. The recommendations give advice on the future direction of the production of the culture codes.

1.5.2 Software and datasets

The main software package used within the research project is *ArcGIS*. This GIS software packages has, on the one hand, been used for visualization purposes (such as in section 4.1). On the other hand, ArcGIS has been used for analysis purposes. It has been used in chapter 5, where the decision rules and decision tree are tested by producing culture codes for the cadastral parcels in a number of test areas in the municipality of Enschede. In chapter 5, an extension for ArcGIS, called *Hawth's Analysis Tools* (SpatialEcology.com 2011), has been used as well. Two tools from this extension, the 'Polygon in Polygon Analysis' tool and the 'Count Points in Polygons' tool, have been used, as these functionalities are not available in ArcGIS. With the 'Polygon in Polygon Analysis' tool "for each polygon in the zonal layer,

statistics are derived from the polygons in the summary layer that overlap that zonal polygon, and the results are written to new fields in the attribute table of the zonal layer" (SpatialEcology.com 2011). With the 'Count Points in Polygon' tool "for each polygon, the number of points that occur inside it are counted and the value is written to the polygon attribute table" (SpatialEcology.com 2011).

Another software package used during the research project is *FME Desktop*. This software package has been used mainly for data preparation purposes, which include data transformations from the MapInfo format to ESRI shapefiles and creating datasets only containing spatial data for the municipality of Enschede. The latter data preparation step is done by making use of the *clipper transformer* (Safe Software 2011a). Figure 1.9 shows how the clipper transformer works. In this research, the transformer was used to create source datasets for the municipality of Enschede from national datasets.

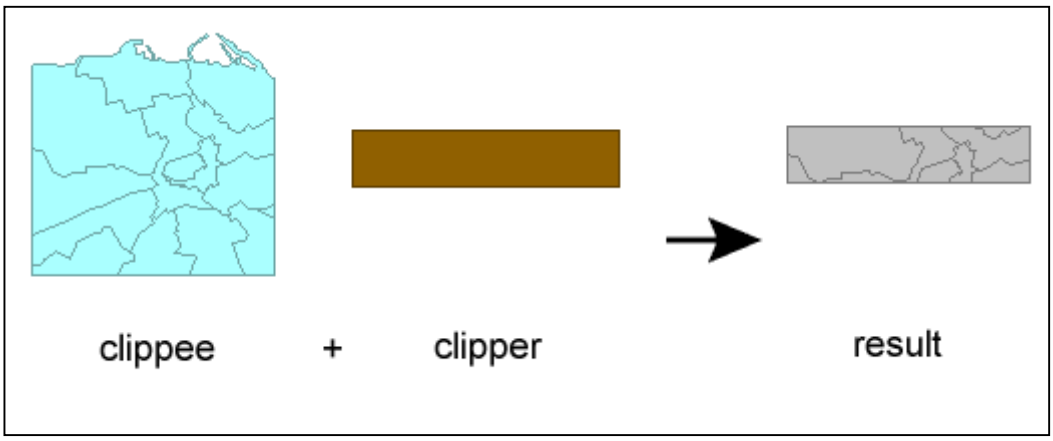


Figure 1.9: Explanation for the clipper transformer in FME Desktop (Safe Software 2011a).

The original plan has been to test the decision rules for the complete municipality of Enschede. However, the large amount of features (e.g. about 57.000 parcels and 150.000 building features) causes serious performance problems. As a result, it has been decided to choose a number of test areas from this municipality. To make sure that the most common situations are present in the testing procedure, four different areas have been chosen:

- A rural area.
- An area in the city centre.
- A living area.
- An industrial/harbor area.

The source datasets used in the research project have all been used in the ESRI shapefile format, as this is the format best supported within the ArcGIS software. Most datasets were obtained from the organization of Kadaster and some of the datasets needed a data format transformation from the MapInfo format to the ESRI shapefile format (GBKN and LKI). The BRP data presented in this research, have been obtained from a WMS (web mapping service) of the dataset, made available through the Internet site http://gisserver1.agro.nl/arcgis/rest/services/BRPGewas2009/MapServer. As it is a WMS, no editing of the data was possible, only presentation. No data were obtained for the LGN dataset, as these data were not available at the organization of Kadaster, and obtaining these data would involve costs (as is explained in section 3.5).

1.6 – Reading guide

This last section of chapter 1 serves as a reading guide for the remaining part of this thesis, as it gives an overview of the contents of the other chapters.
In chapter 2 theoretical background is provided on the concepts central to the topic of this thesis. The chapter starts by explaining the concepts of schema integration and schema mapping (section 2.1). Afterwards, in section 2.2, issues and problems that arise when using different spatial datasets are described. This includes a general introduction on spatial data(sets), spatial data infrastructures (SDIs), schema modeling and schema modeling languages, and semantic, geometric and temporal issues when using different spatial datasets for the production of "new" data or datasets.

Then, in chapter 3, the candidate source datasets listed in section 1.5.1 are described and analyzed: BAG (section 3.1), BRP (3.2), TOP10NL/BRT (3.3), GBKN/BGT (3.4), LGN (3.5) and AKR/LKI (3.6). First, for all of these dataset a general introduction is given on the contents and production. Second, the useful classes and attributes for the production of the culture code dataset are listed, in combination with their geometry types. Third, visual mappings are created between these useful classes and/or attributes and the culture codes. Fourth, the quality and up-to-dateness of the datasets is analyzed. Fifth, and finally, the strengths and weaknesses of the datasets are listed. The final section of chapter 3 (section 3.7) deals with the decision on which source datasets are used in the remaining part of the research.

In chapter 4, the possibilities of using the selected source datasets are explored. Section 4.1 presents the possibilities and problems that arise when overlaying different source datasets and the cadastral parcels dataset (LKI). Then, in section 4.2, decision rules are formulated which should help to produce the culture codes. This includes decision rules on deciding whether a cadastral should get a non-built-up or built-up culture code or a combination. Furthermore, it includes decision rules which should help to decide on which built-up or non-built-up culture code should be given to a cadastral parcel. Finally, a decision rule describes when the culture code "Special properties" should be given to a cadastral parcel. These decision rules are integrated into a decision tree in section 4.3. Section 4.4 discusses to what extent the decision rules and decision tree cover the culture codes listed in appendix A. Can all culture codes be produced by applying these decision rules and this decision tree? Or are there any culture codes that cannot be produced with this method?

The decision rules and decision tree from chapter 4 are tested in the proof of concept in chapter 5. This proof of concept consists of two parts: a case study and a quality analysis. Case studies and quality analyses are applied to all four test areas: a rural area (section 5.1), a city centre area (5.2), a living area (5.3) and an industrial/harbor area (5.4). In section 5.5 the results of these case studies and quality analyses are summarized and analyzed.

Finally, in chapter 6, the conclusions and recommendations are formulated. The conclusions will answer the main research questions formulated in section 1.4. The recommendations serve as input for further research that could be carried out in order to improve the method proposed in this thesis and/or to automate the procedure for the production of the culture codes.

Chapter 2 – Theoretical background

In the previous chapter background information has been given on the concepts of land registration and cadastre, the culture codes and the reasons why the Dutch institution Kadaster is interested in a new way of producing these culture codes. The current chapters presents a theoretical overview of a number of themes relevant to the research project described in this document. First, a short introduction into the concepts of schema mapping and schema integration is provided (section 2.1). Second, it is discussed what is needed to use, and what the problems are when using, multiple different (spatial) datasets (section 2.2). Finally, a more detailed description is provided on data quality issues (section 2.3). Together, the themes presented in this chapter give direction to the research as described in this document.

2.1 – Schema mapping and integration

As described in the introduction of this document, the institution of Kadaster wants to define an automated way of producing the culture code dataset. This automated production should replace the current ad-hoc, manual production of the culture codes. The possibility for automatically creating the culture codes is offered by schema mapping or schema integration. Although schema mapping or integration is not technically implemented as part of the research project presented in this document, a short introduction into the concepts is included in this document, as the concepts are expected to have a significant role in the following-up project and as the current research project includes a visual mapping between the relevant classes of the source datasets and those of the culture codes dataset.

Mapping between classes or features from source datasets and a target dataset is often referred to as *schema mapping* (Alexe et al. 2008; Batini, Lenzerini, and Navathe 1986; Fagin et al. 2009). Alexe et al. (2008, p. 10) describe schema mapping as "the specification of the relationships between a source schema and a target schema". An example of a mapping scenario is presented in figure 2.1. The source database ('CompDB') contains a set of projects as well as a set of employees serving as supervisors in the projects, both sets containing a number of attribute classes. The target dataset ('OrgDB') combines information from the sets of projects and employees from the source dataset into one new set of projects. The information which should be available in the target dataset is the projectname ('pname' from the set of projects) and the name and contact information of the supervisor ('ename' and 'contact' from the set of employees). The arrows in the figure visualize the mappings.

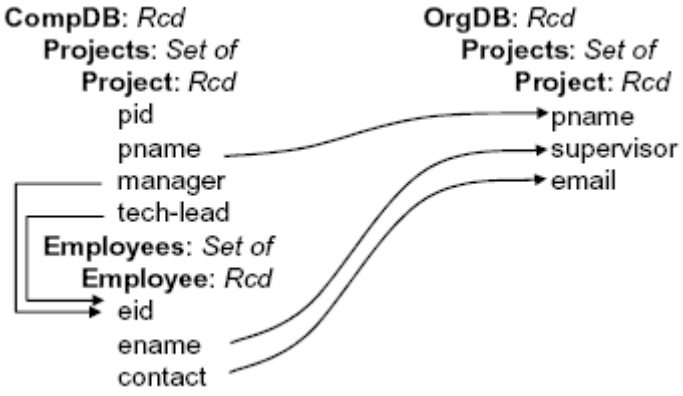


Figure 2.1: Mapping scenario (Alexe et al. 2008, p. 8).

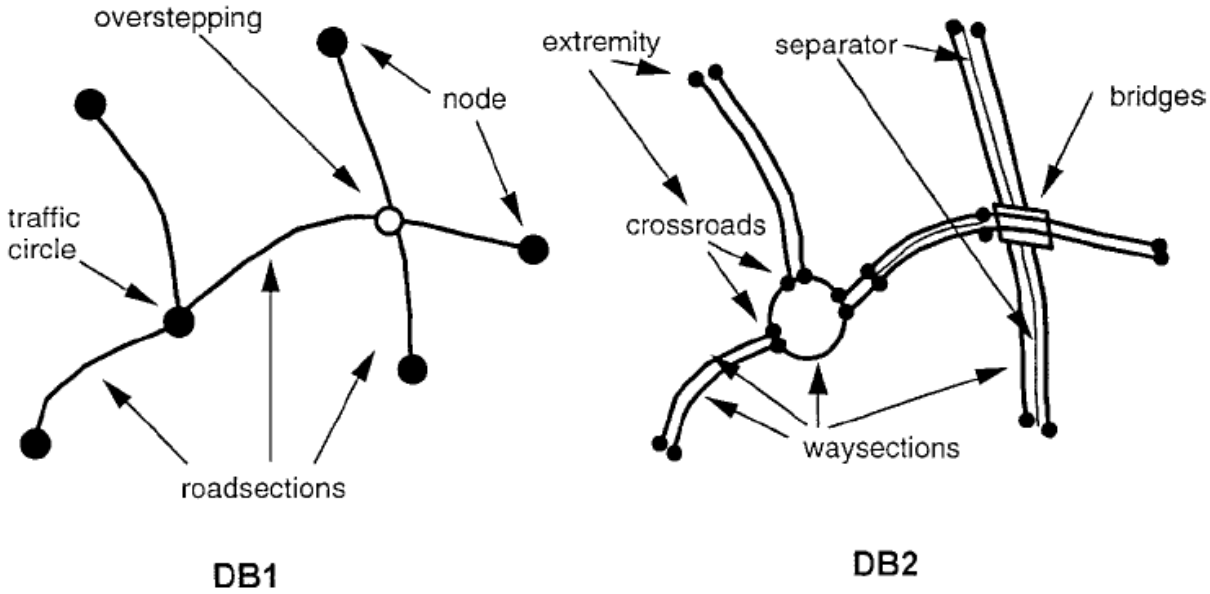


Figure 2.2: Two databases (DB1 and DB2) with different viewpoints on the same road network (Devogele et al. 1998, p. 345).

Another, closely related, way of obtaining and combining information from source datasets into a new target dataset is *schema integration*. Devogele et al. (1998) define an integrated schema as an important component of database integration. They describe database integration as taking "as input a set of databases (schemas and data instances), and [producing] as output a single unified description of the input schemas (called the integrated schema) and the associated mapping information supporting integrated access to existing data instances through the integrated schema" (p. 336).

An example of schema integration is provided by Devogele et al. (1998). In the example, two databases (DB1 and DB2) have to be integrated into one new schema. Both datasets contain information on a road network (see figure 2.2), but with different description methods. The UML class diagrams of the source datasets are presented in figures 2.3 and 2.4. A schema integration procedure implies that the different classes distinguished between in the source datasets are combined in the target dataset. Figure 2.5 shows the integrated schema, in which a road section (from DB1) is an aggregation of way sections and separators (from DB2). Furthermore, a node (from DB1) is an aggregation of way sections and extremities (from DB2), where an extremity is "a point or an area where traffic conditions change", such as crossroads, tolls or ends of tunnels (Devogele et al. 1998, p. 341). Finally, an upper-class 'Bridges/Oversteppings' has been added, being an aggregation of oversteppings from DB1 and bridges from DB2.

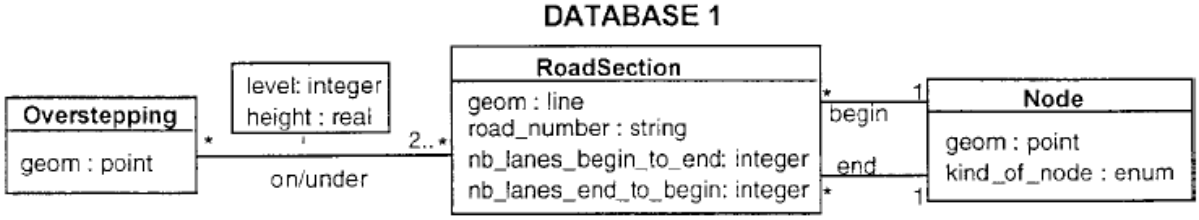


Figure 2.3: UML class diagram of DB1 (Devogele et al. 1998, p. 339).

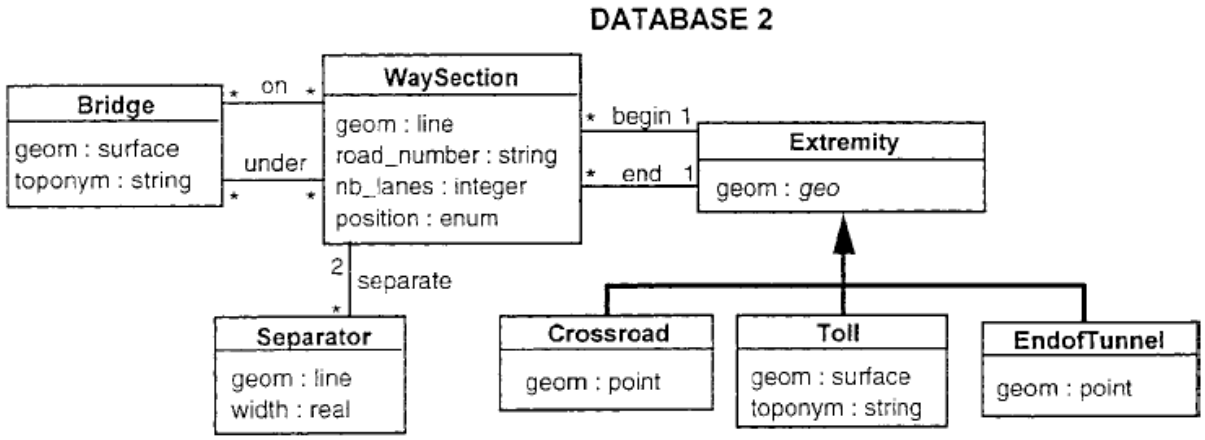


Figure 2.4: UML class diagram of DB2 (Devogele et al. 1998, p. 340).

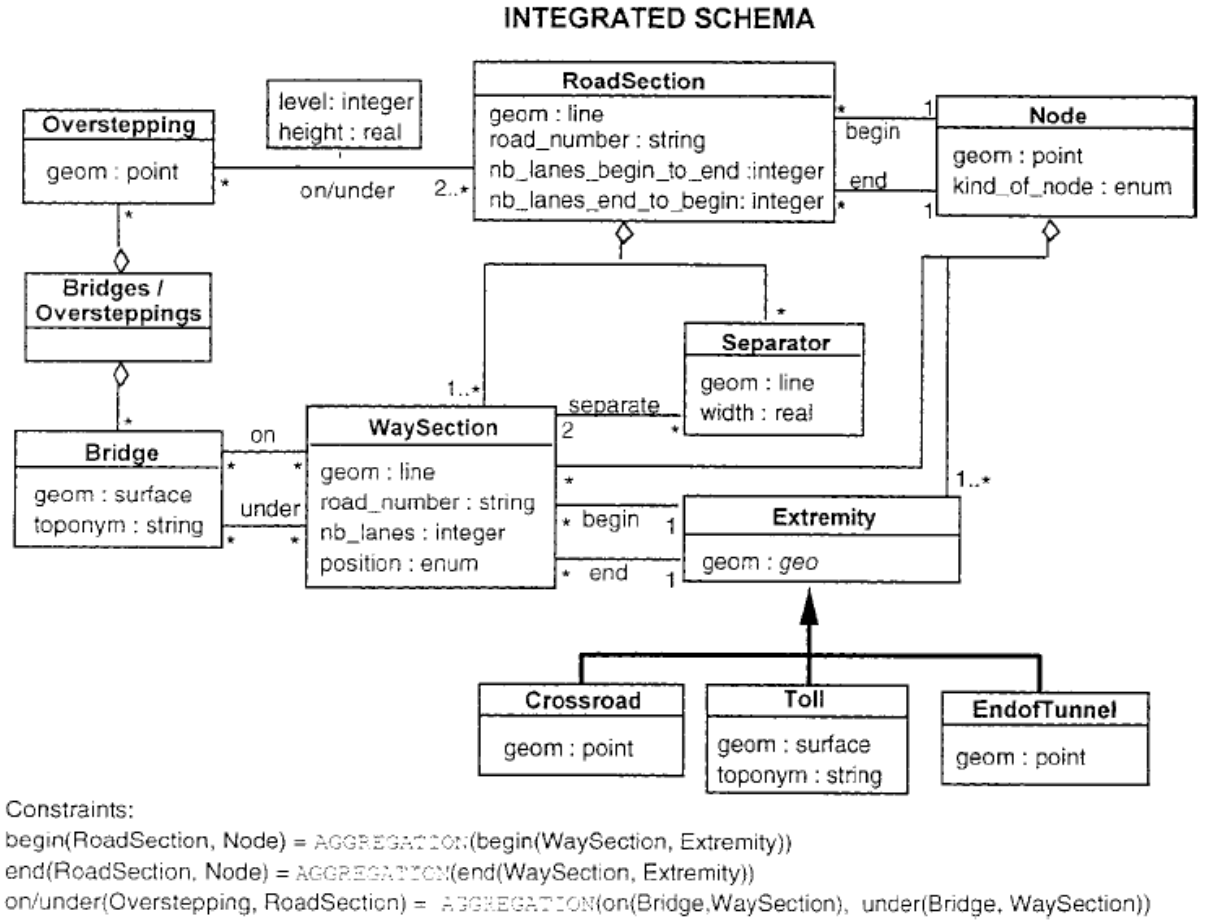


Figure 2.5: Integrated schema for source datasets DB1 and DB2 (Devogele et al. 1998, p. 348).

Summing up, there are two closely related ways of combining information from source datasets into one new target database: schema mapping and schema integration. The main difference between these two methods is, that in the case of schema integration all classes from the schemas of the source datasets are copied and integrated into the target dataset's schema, whereas in the case of schema mapping only the information is collected which is relevant to the purpose of the target dataset. In the latter case, therefore, only a number of classes from the source dataset's schema are used to obtain data for the target dataset.

2.2 – Using different spatial datasets: issues and problems

Reusing (geographical) information from different datasets is not without difficulties. Therefore, this section presents the main points to be taken into account when performing data integration with different source datasets. As there are some special issues involved when dealing with geographical information, a short introduction into the concept of spatial data/geographical information is a useful start of the chapter (section 2.2.1). Second, theoretical background is provided on the concept of SDI (spatial data infrastructure), as it is an important concept for re-using and sharing geographical information (section 2.2.2). Section 2.2.3 gives a basic introduction into the concept of schema modeling, by presenting two languages for schema modeling: OWL and UML. Finally, section 2.2.4 describes the main issues to be considered when using different (geographical) datasets: semantics and geometry. Spatial data quality issues, including temporal characteristics of datasets, are described in a separate section in this chapter (section 2.3).

2.2.1 Spatial data(sets)

A *spatial dataset* is a collection of spatial data. But, what is exactly meant by the term "spatial data"? According to Heywood et al. (2006, p. 32), "data are observations we make from monitoring the real world". What makes spatial data special is the linkage between the data and a specific location. A term which is often interchangeably used is *geographic information* (or: geo-information), which is defined by the Dutch normalization institution NEN as data with a direct or indirect link to a location on the Earth's surface (NEN 2005, p. 13) and by Goodchild (2000, p. 345) as "information about the distribution of phenomena on the surface of the Earth (and the near-surface)". Goodchild also gives a more formal definition of geographic information: "a well-defined subset of information in general, and […] a commodity that is independent of the media on which it is stored, communicated, and used, and of the structures and models used to represent it" (Goodchild 2000, p. 345). Finally, Van Loenen et al. (2008, p. 11) explain that geographic information links together location, time and attributes. This last description of geographic information resembles the three modes or dimensions of data distinguished by Heywood et al. (2006, pp. 32-34): temporal, thematic and spatial. Van Loenen et al. (2008, p. 11) also stress that one can distinguish between tangible and non-tangible geographic information. Examples of tangible geographic information are houses, roads and trees. An example of non-tangible geographic information is a municipal boundary, which may not be recognizable in the landscape. As Heywood et al. (2006, p. 32) argue that one needs to add meaning or context to data in order to let them become information, in this document the term spatial data is used.

2.2.2 Spatial data infrastructures (SDIs)

As the production of spatial data and spatial datasets is a very expensive activity (van Loenen et al.2008), selling, sharing and the re-use of the datasets is very important. In order to facilitate the accessibility, exchange/sharing and use of spatial data the concept of SDI (*spatial data infrastructure*) has been introduced (Crompvoets 2009; de Man 2007; Feeney and Williamson 2000; Rajabifard and Williamson 2001). Five components (visually presented in figure 2.1) building an SDI can be distinguished:

1. **People**: this includes SDI policy-makers, data suppliers, as well as those using the data made available through the SDI;
2. **Technology**: this includes hardware and software used for the SDI, but above all those technologies used to share data, such as WMS (web map service) and WFS (web feature service);

3. **Policies**: this is about the policies regarding the architecture of the SDI itself, as well as regulations with regard to access to spatial data and pricing;
4. **Standards**: standards for the SDI, the datasets and the metadata are crucial with regard to the sharing of data through the SDI;
5. **Spatial data**: the spatial data which are to be shared and used through the SDI.

(Crompvoets 2009; Rajabifard and Williamson 2001)

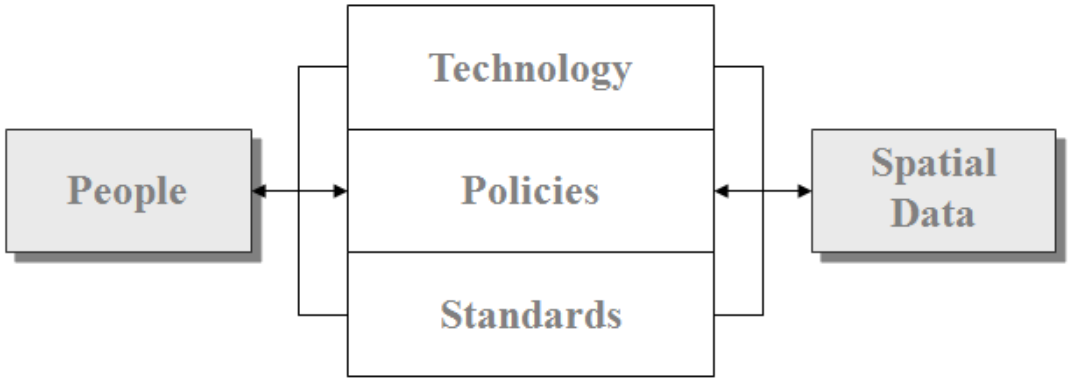


Figure 2.6: The five components building an SDI (Crompvoets 2009).

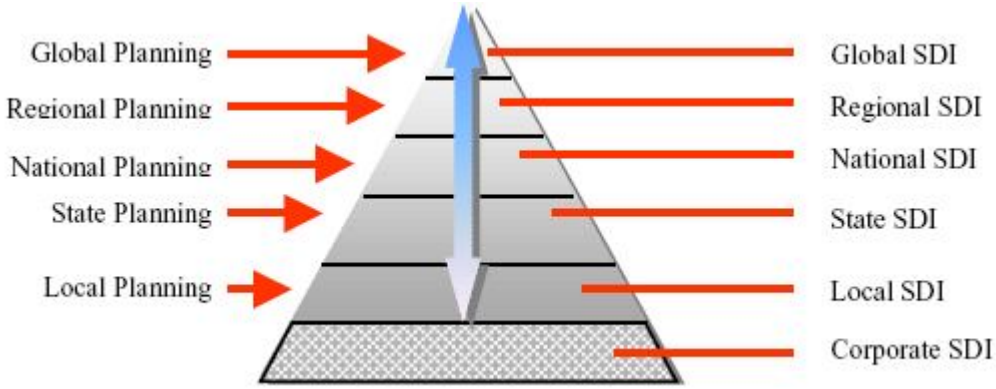


Figure 2.7: The SDI hierarchy (Rajabifard et al. 2000, p. 7).

Figure 2.7 shows that SDIs exist at different hierarchical levels: global, regional, national, state, local and corporate level. Different actors at these levels share and re-use spatial data. However, at the same time, sharing and re-using of spatial data takes places between actors of different hierarchical levels (Rajabifard et al. 2000), visualized by the blue/white arrow in the middle of the figure. For example, actors at a regional level may use spatial data from an actor at the local level and vice versa.

Spatial data infrastructures and their components facilitate easier re-using and sharing of spatial data, and are therefore useful for the production of the culture codes, as this production should rely on existing spatial datasets. The components 'Technology' and 'Standards' are of main importance for the technical production of the culture codes by means of schema mapping or integration. As this research document mainly focuses on the contents and characteristics of a number of potential source datasets in order to define the possibilities and difficulties when using them for the production of the culture codes, the component 'Spatial data' is the most important SDI component for this research project. Furthermore, the research project presented in this document takes into account the component of 'Policies', as it is described whether the institution of Kadaster has access to the data, and whether this access involves a fee to be paid.

2.2.3 Schema modeling

Datasets are always representations of real-world phenomena; how these phenomena are represented in a dataset is described by (schema) models. Miller (2003) states that a "model abstracts the essential details of the underlying problem from its usually complicated real world". As such, schema models should explain how a dataset describes part of the real world. Modeling of schemas can be done in a number of ways. Two of the most commonly used modeling languages are OWL and UML class diagrams. A very basic introduction into these two modeling languages is provided in this section.

OWL (Web Ontology Language) 'is designed for use by applications that need to process the content of information instead of just presenting information to humans' (W3C 2004). Furthermore, OWL contains 'vocabulary for describing properties and classes: among others, relations between classes (e.g. disjointness), cardinality (e.g. "exactly one"), equality, richer typing of properties, characteristics of properties (e.g. symmetry), and enumerated classes' (W3C 2004). UML (Unified Modeling Language) class diagrams give "an overview of a system by showing its classes and the relationships among them" (Miller 2003).

Although both OWL and UML are describing classes and the relationships among these classes, there are profound differences between the two modeling languages. Xu et al. (2008) describe these differences and stress that the two languages have been developed for different purposes: 'UML has a visual design notation, which is much more human-readable and OWL is derived from Descriptive Logics, which is meant for inference' (p. 2). OWL mainly focuses on the ontology – described by Xu et al. (2008) as the characteristics/specifications of a certain conceptualization – of a certain system. Figures 2.8 and 2.9 show examples of UML and OWL. The figures, both describing generalization within a certain model, clearly show that UML is visually oriented, whereas OWL is a descriptive representation of the model. Both UML and OWL use a hierarchical way of structuring the schema (e.g. in figure 2.9, municipalities and police are both 'below' the class sector).

The purpose of this section has not been to fully explain the concepts of OWL and UML. Rather, the purpose has been to give an brief introduction into schema modeling and two modeling languages.

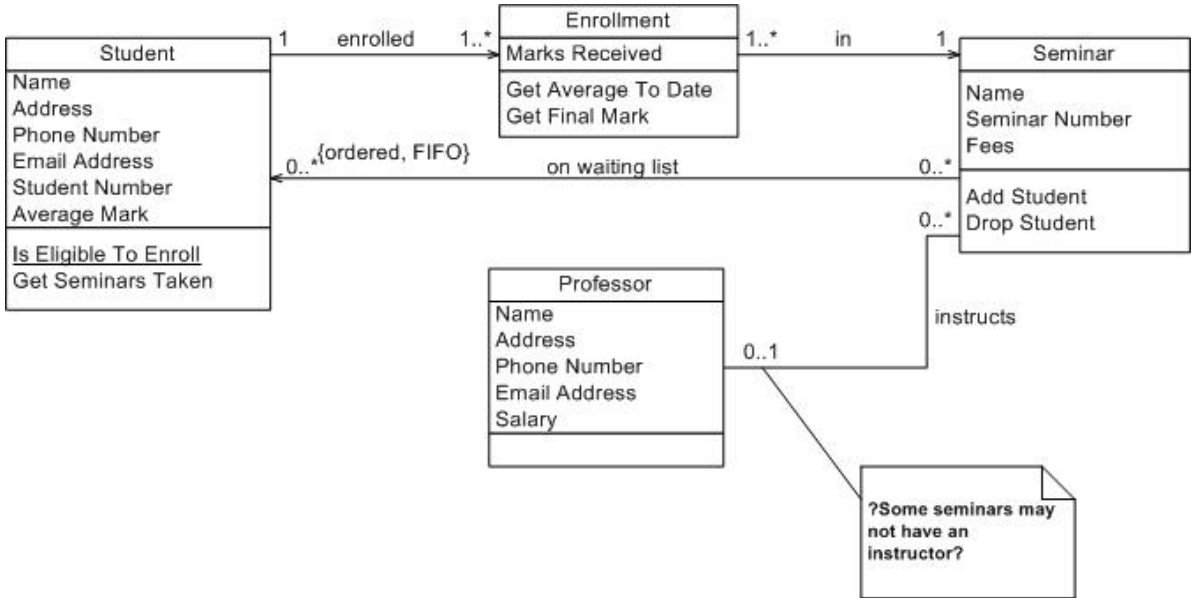


Figure 2.8: A classical example of a UML class diagram (Ambler 2010).

```
<owl:Class rdf:ID="FireBrigade">
  <rdfs:subClassOf rdf:resource="#Sector"/>
</owl:Class>
<owl:Class rdf:ID="MedicalService">
  <rdfs:subClassOf rdf:resource="#Sector"/>
</owl:Class>
<owl:Class rdf:ID="Municipalicity">
  <rdfs:subClassOf rdf:resource="#Sector"/>
<owl:Class rdf:ID="Police">
  <rdfs:subClassOf rdf:resource="#Sector"/>
</owl:Class>
```

Figure 2.9: OWL example (Xu et al. 2008, p. 7).

2.2.4 Issues when using different datasets: semantic, geometric and temporal

As referred to in section 2.2.2, producing spatial data is very expensive, and therefore the re-use of spatial data is attractive and important. However, reusing spatial data is not without difficulties. Devogele et al. (1998, p. 335) formulate the difficulties when re-using existing data as follows: "reuse for new applications is a nightmare, due to poor documentation, obscure semantics of data, diversity of data sets (what information is stored, how it is represented and structured, what quality it has, which date it refers to, which scale is used, …), and the heterogeneity of existing systems in terms of data modeling concepts, data encoding techniques, storage structures, access functionalities, etc." This citation makes clear there are many obstacles to overcome when re-using existing spatial data(sets). This section explores the, in relation to the production of the culture codes, most important obstacles: issues related to semantics, geometry and data history. Quality-related issues, another important set of issues for the research described in this document, are to be described in section 2.3.

Terms used in (spatial) datasets can refer to different meanings and can be interpreted differently by various individuals. Semantics should help to make clear what is meant by the terms used by producers of the dataset and should, as a result, help to avoid confusion and misinterpretation. Furthermore, semantic similarity – when similar terms used in different datasets refer to the same meaning – is an important tool for reusing spatial data. According to Schwering (2008, p. 5) "semantic similarity is central for the functioning of semantically enabled processing of geospatial data". The author continues by arguing that, among others, similarity "is the basis for semantic information retrieval and integration" (Schwering 2008, p. 5). As semantic information retrieval and/or integration are central to the schema mapping or integration procedures to be carried out in order to produce the culture codes, semantic similarity is an important concept during the project. Some of the issues related to semantic similarity are:

- Similarity in classification between datasets. This means that classes in different datasets have a similar content and definition. In the case of a mapping between a number of source datasets and a target dataset, it implies that the classes defined in the target dataset have a relatively large similarity with a number of classes from the source datasets. For example, if both the TOP10NL source dataset and culture code target dataset contain an agricultural land use classification, similarity in classification would mean that both datasets distinguish the same classes (e.g. cropland, grassland and cultivation).
- Similarity in definitions. This means that terms have the same meaning when used in different spatial datasets. For example, if the term "roads" is used in the TOP10NL

source dataset and the culture code target dataset, in one dataset it could include only main roads, whereas in the other it could include all kinds of different road types (e.g. walking paths, cycling paths, local roads, paved and unpaved roads, and freeways).

In this thesis, another semantic issue plays an important role: translation. The culture code dataset and the candidate source datasets are Dutch datasets using Dutch terms. For the purpose of this thesis these terms have been translated into English. However, this is not without risks, as the translation could be wrong or the English term might not completely cover the meaning of the Dutch term. Therefore, in this thesis the Dutch names for object classes and attributes have been included (sometimes in the main text, such as in section 3.5, and sometimes in the appendix).

When integrating different *spatial* datasets, next to semantics, geometric aspects of the datasets and its classes and instances play an important role. First, real-world artefacts or phenomena can be represented in different ways in the dataset. Kraak & Ormeling (2003, p. 3) state that, for geospatial or geometric data, one can differ between "point-, line, area- or volumetrically shaped objects". The chosen representation depends on issues such as scale and level of generalization, as well as on the phenomenon represented by it. The type of representation used in the datasets strongly influences the way it can be used for the production of a new dataset. For example, with area-shaped objects it is possible to calculate percentages covered by a certain class or attribute, whereas with point- or line-shaped objects such calculations are not possible.

Second, the geometry and geometry type of the objects in the target dataset may differ from those in the source datasets. In relation to the research presented in this document, the culture code dataset consists of area-shaped objects only, whereas the source datasets may consist of a combination of point-, line- and area-shaped objects. Furthermore, the area-shaped objects from the target dataset may refer to other geographical areas than the area-shaped objects from the source dataset. For example, a parcel may – and, actually, does in many cases – not completely coincide with a land use area from one of the source datasets.

Third, when using different source datasets, geometric objects from one of these datasets may overlap with geometric objects from another one. This is no problem, as long as the target dataset allows a single location to have several land use types, artefacts or phenomena located on it. It is a problem, however, when the target dataset does not allow this. This implies that in the case of the culture code dataset, which allows for only one built-up and one non-built-up code per parcel, overlapping features from the source datasets could cause problems in defining the right culture code and it might prove to be necessary to define generalization decisions (Foerster et al. 2010). These decisions include decisions on the priority given to the different datasets.

Next to the semantic and geometric issues, temporal issues play an important role in the production of the culture code dataset. According to Worboys (1994, p.26) "much information which is referenced to space is also referenced to time". Figure 2.10 visualizes how spatial data can change over time. In the case of the culture codes, the historical changes made to the spatial data are of importance, as it should be possible to calculate the right culture code for every moment in the past. This means that the candidate source datasets should contain historical data to facilitate such historical calculations.

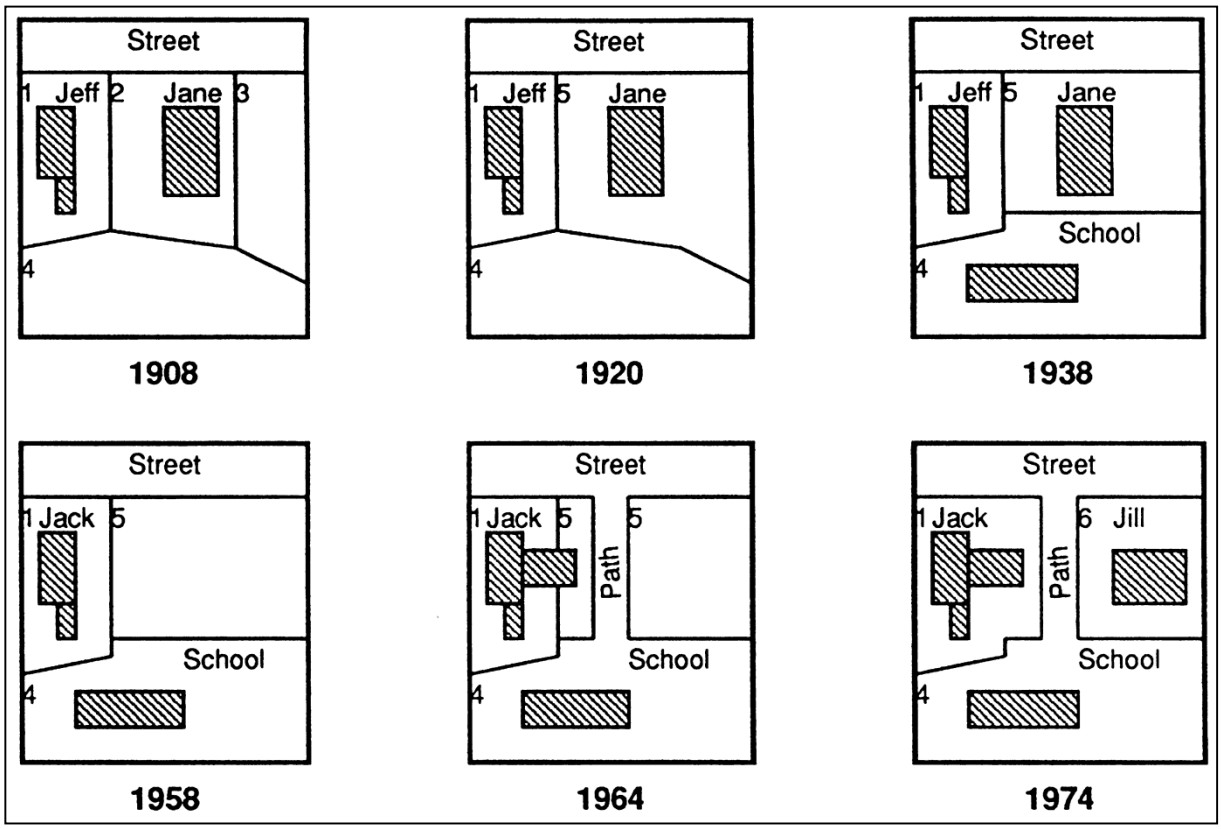


Figure 2.10: Spatiotemporal variation of land ownership (Worboys 1994, p. 28).

2.3 – Spatial data quality issues

The quality of 'new' data created by integrating information from different source datasets depends largely on the quality of these source datasets. This section, therefore, introduces the basics of spatial data quality. By referring to an unpublished document written by Crain in 1990[4], Morrison (1995, pp. 8-11) defines a list of elements of spatial data quality. These elements are:

- *Lineage*. Van Oort (2006, p. 14) defines lineage as recording the "the history of a geographic data set. A description of the source material from which the data were derived, and the methods of derivation, including all transformations involved in the production process." Clarke & Clark (1995, p. 13) would agree on this definition, as they also define lineage as the history of a dataset. Furthermore, they emphasize that a spatial dataset is not a static entity, but that it is rather dynamic. Lineage makes sure that the changes are recorded, helps the user of the dataset to find information on the changes made to that particular dataset, and, as such, gives the dataset a higher level of quality (Stuiver 2010).
- *Positional accuracy*. "In the mapping sciences the position of a real world entity is described by values in an appropriate coordinate system. Positional accuracy represents the nearness of those values to the entity's "true" position in that system." (Drummond 1995, p. 32). In other words, positional accuracy calculates the difference between a feature's location on the map and the location of that same feature in the real world. The positional accuracy is often defined by calculating the Root Mean

[4] The title of the unpublished work by Crain is "Classification system for ICA digital data quality assessment" of the ICA Commission on Spatial Data Quality.

Square Error (RMSE) or standard deviation (see Drummond 1995, pp. 36-40). Van Oort (2006, p. 15) argues that this type of positional accuracy should be referred to as "absolute positional accuracy" and that there is a difference between absolute and relative positional accuracy. According to Van Oort (2006, p. 15) "relative positional accuracy is the accuracy relative to other data in the same test data set".

- *Attribute accuracy*. According to Goodchild (1995, p. 60) a "host of different kinds of uncertainty affect attributes, particularly when the process of assigning attributes to locations or features is long and complex, as is often the case". Examples of such uncertainty are the inaccuracies of measuring instruments, and subjectivity when assigning locations to a particular land use class. Van Oort provides an excluding definition by stating that "attribute accuracy is the accuracy of all attributes other than the positional and temporal attributes of a spatial data set" (van Oort 2006, p. 15).
- *Completeness*. Brassel et al. (1995, p. 87) define completeness as a spatial data quality criterium that "indicates whether each entity instance is present and whether all of its attributes are present, where the totality of entity instances is defined by the entities within the abstract universe". A different definition of completeness is provided by Van Oort (2006, p. 15): "completeness is a measure of the absence of data and the presence of excess data". This is a broader definition than the one provided by Brassel et al., as it also includes the presence of excess data. Van Oort (2006, p. 15) uses the term errors of commission for cases of overcompleteness, whereas he uses the term errors of omission for cases of incompleteness.
- *Logical consistency*. According to Kainz (1995, p. 109) "logical consistency deals with the logical rules of structure and attribute rules for spatial data and describes the compatibility of a datum with other data in a data set". Van Oort (2006, p. 15) argues that Kainz focuses too much only on topological consistency in his description of logical consistency, emphasizing too little on other aspects of logical consistency, such as: valid values, graphic data, date, geometric consistency, semantic consistency, conceptual consistency and format consistency.
- *Semantic accuracy*. Salgé (1995, p. 139) defines semantic accuracy as a spatial data quality criterium referring "to the quality with which geographical objects are described in accordance with the selected model".
- *Temporal information/quality*. Guptill (1995, p. 153) describes the importance of temporal information as follows: "Information such as when a data element was corrected or revised is an important factor in judging data quality. Although users often want the most current information, historical information (or information on just the changes over time) is necessary for any process studies." Guptill distinguishes between three types of temporal information (or: times):
    - Logical time or event time: the time at which the change actually occurred;
    - Observation time or evidence time: the time at which the event was observed;
    - Transaction time, database time or capture time: the time at which the event was added to the database." (Guptill 1995, p. 155)

  Van Oort (2006, p. 17) uses the term temporal quality, instead of temporal information, and presents a list of indicators for defining the temporal quality of a spatial dataset. These indicators and a short description can be found in table 2.1.

Indicator of temporal quality	Description of the indicator
1. Accuracy of time measurements	Summary of errors in time measurements.
2. Temporal validity	The validity in respect of time, also sometimes called currency. According to CEN the temporal validity can take on one of the following three values: “out_of_date”, “valid” or “not_yet_valid”.
3. Temporal consistency	Correctness of the order of events.
4. Last update	Last time the data were updated.
5. Rate of change	An estimate of the rate of change in the phenomenon represented in the data. Together with information on the last update this element can inform the user about the currency.
6. Temporal lapse	The average time between change on the nominal ground and its representation in data.

Table 2.1: Indicators of temporal quality and a description of these indicators (taken fromvan Oort 2006, p. 17).

Next to the seven spatial data quality criteria described above, Van Oort (2006) defines four more elements of spatial data quality. These elements are:

- *Usage, purpose, constraints*. Van Oort (2006) describes that for a spatial dataset “intended use (purpose) is not necessarily the same as actual use (usage)”. Furthermore, a number of constraints can be distinguished: direct costs (what has to be paid by a user to obtain and use the data), indirect costs (“the time and material used to make the data ready for use for the buyer”) and “legal or contractual constraints to the access and application of data” (van Oort 2006, pp. 16-17).
- *Variation in quality*. According to Van Oort (2006, p. 17) variation in quality describes the differences in quality within the dataset.
- *Meta-quality*. Van Oort (2006, pp. 17-18) describes meta-quality as an element that “provides information on the quality of the quality description. For example if the positional accuracy is estimated from a smaller sample size, then that estimate is of lower quality.”
- *Resolution*. Resolution is described by Drummond (1995, p. 36) as indicating “the smallest measurement possible by the instrument (e.g. digitizer, scanner) system”, but the author does not see it as an extra spatial data quality criterium. Van Oort (2006, p. 18), unlike Drummond, does see it as an extra spatial data quality criterium, justifying it by stating that “often a decision or analysis requires data at a certain resolution and as such, information on the resolution is important in the first step of fitness-for-use assessment”.

Data quality elements are also specified in an ISO standard: “ISO 19115:2003 – Geographic information – Metadata” (ISO 2011). Figure 2.11 shows that the ISO standard on metadata distinguishes between five data quality elements:

- Completeness;
- Logical consistency;
- Thematic accuracy;
- Temporal accuracy;
- Positional accuracy.

(Van Oosterom et al. 2011)

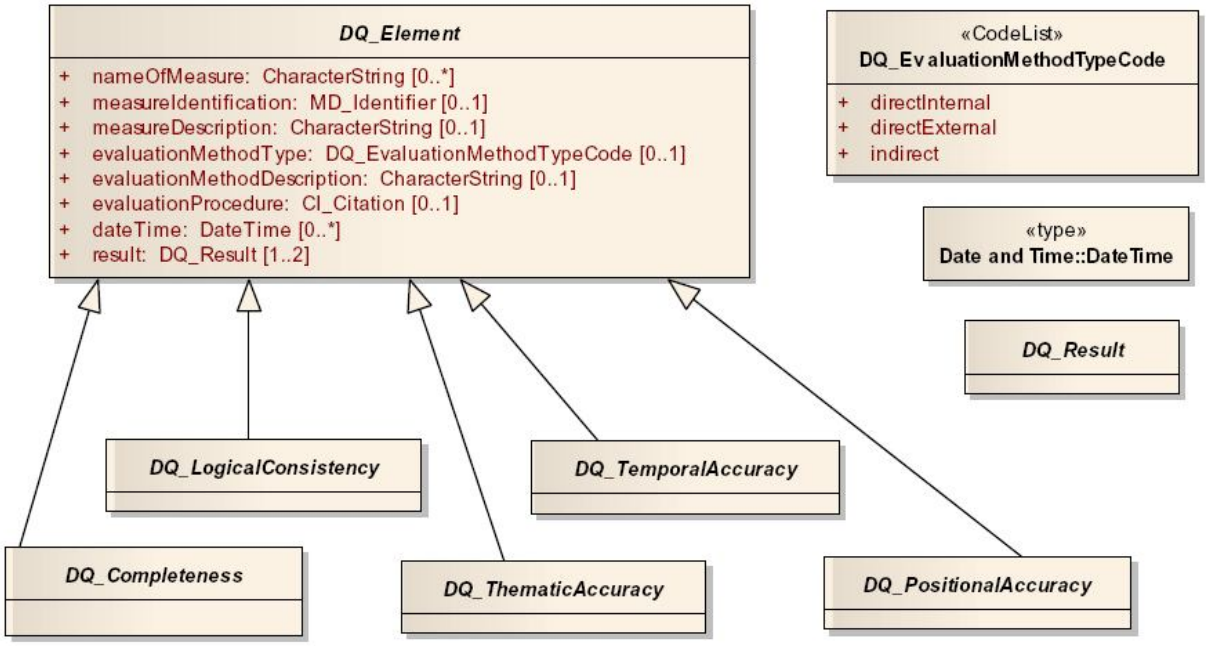


Figure 2.11: UML class diagram of data quality elements specified in ISO 19115:2003 (Van Oosterom et al. 2011).

Completeness, logical consistency and positional accuracy are also present in the list of eleven spatial data quality elements presented above. The other two elements from the ISO standard, thematic accuracy and temporal accuracy, are also presented there, but have slightly different names. In the list presented above, for thematic accuracy the name semantic accuracy is used, and for temporal accuracy the name temporal information/quality. Lineage is also included in the ISO 19115:2003 standard. However, it is not seen as a subclass of DQ_Element – the superclass in figure 2.11. Rather, it is presented as part of the core metadata for geographic datasets (Geonovum 2009).

Chapter 3 – Candidate source datasets

For the production of the culture codes, a number of candidate source datasets have been listed in the introduction of this document. For these six datasets – BAG, the Dutch key register of agricultural parcels of the Ministry of Agriculture (BRP)[5], TOP10NL/BRT, GBKN/BGT, LGN and AKR/LKI – detailed meta-information is provided in this chapter. The information has been obtained by using catalogues and other descriptive literature about the datasets. Furthermore, for all datasets, except the LGN, interviews have been performed with specialists. The interviews are another type of source – next to the catalogues and literature sources – and are, therefore, treated as such in this document. The specialists interviewed are:

- Frank Kooij – BAG specialist working for Kadaster
- Wim Hendriks & Janette Dijk – Key register of agricultural parcels specialist for the Ministry of Economy, Agriculture and Innovation
- Daniël te Winkel – TOP10NL and BRT specialist working for Kadaster;
- Bart Maessen – GBKN and BGT specialist for Kadaster.

For all datasets a short introduction on the general content and production method is followed by an overview of the useful classes and their geometry type. Then, conceptual mappings are presented showing the possible links between the attributes of the candidate source datasets and the culture codes. Furthermore, the quality and up-to-dateness of the candidate source datasets is discussed. Finally, the strengths and weaknesses of the datasets are presented.

The chapter consists of sections on the BAG (3.1), BRP (3.2), TOP10NL/BRT (3.3), GBKN/BGT (3.4) and LGN (3.5) and AKR/LKI (3.6). In section 3.7 a decision is made on which candidate source datasets are used in the remaining part of the research presented in this thesis.

3.1 – BAG

3.1.1 Introduction

The BAG dataset, visualized in figure 3.1, is one of the key registers defined by the Dutch government. The Dutch government has introduced these key registers in order to guarantee a certain content, up-to-dateness and level of quality, independent from those different parties involved with building and filling the datasets (E-overheid 2009). As a result, the key registers seem reliable datasets to be used in the creation of the culture code dataset.

The BAG (key register of addresses and buildings) consists of two parts: (1) BRA, the key register of addresses, and (2) BGR, the key buildings registration (Fuld and Rietdijk 2004, p. 3). The BRA contains number denotations, public spaces and domiciles. The BGR contains information on mooring sites ('*ligplaats*'), buildings, footholds ('*standplaats*') and residential objects '*verblijfsobjecten*' (Ministerie van VROM 2009a). For the purpose of creating the culture codes dataset, the BGR is most useful and is described in more detail in this section. In the remaining part of this document the source dataset will be referred to as the "BAG

[5] The abbreviation BRP stands for "Basisregistratie Percelen", which means key register of parcels. However, as in the BRP dataset parcels refer to agricultural parcels (parcels used for a certain type of agricultural activity) and not to cadastral parcels, it is chosen to use "key register of agricultural parcels" in this thesis when referring to the BRP dataset. In this thesis, "parcels" is sometimes used for referring to cadastral parcels.

dataset", as the name "BAG" is more commonly used in Dutch governmental and scientific publications, although in practice only the BGR is used for the project.
The dataset is produced and maintained by the Dutch municipalities. Every municipality has its own registration which is copied to a national central database, the LV[6]. The municipalities deliver the attribute values. In order to guarantee the correctness of the data in the BAG dataset, for every attribute value a source document is needed (Ministerie van VROM 2009c). These source documents contain information on the reason why certain data are added to the BAG dataset. Examples of such source documents are building permissions and documents describing the creation of a new address. The spatial data are produced by land surveying methods. Some municipalities are themselves performing the land survey, others leave the land surveying to the Stichting GBKN (see section 3.4 for more information).

Furthermore, the geometry of the BAG buildings refers to the view from above. This means that not always the geometry of the building in the BAG dataset refers to the real-world geometry on ground-level. Finally, the BAG dataset contains historical data, as former buildings and residential objects are maintained in the databases. The status of the buildings and residential objects shows whether they are still in use or not.

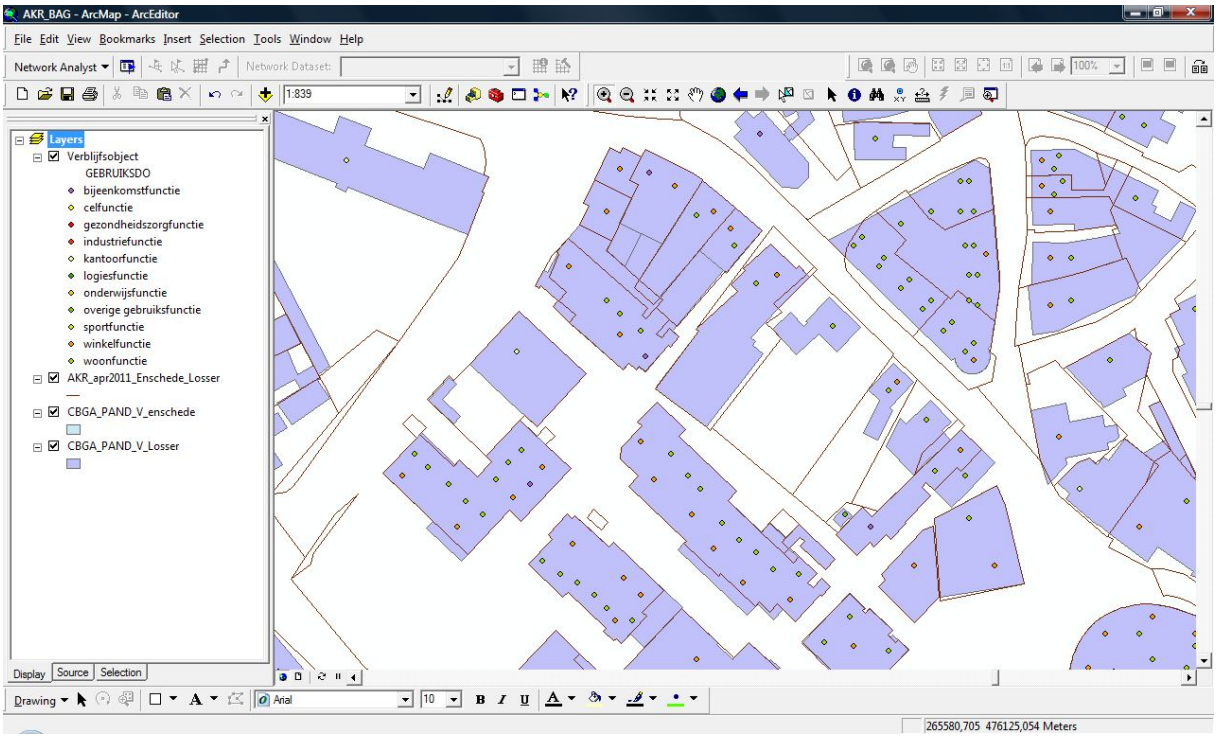


Figure 3.1: Screenshot of the BAG dataset in ArcGIS. The blue area-shaped objects represent buildings and the point-shaped objects represent the verblijfsobjecten.

3.1.2 Useful classes/attributes and geometry type

Table 3.1 shows the useful object classes and attributes of the BAG dataset. It also contains information on the geometry type of the object classes, although in the BAG dataset the geometry of the features one of the attributes distinguished between (Ministerie van VROM 2009b, p. 7). It should be mentioned here that the residential object points are always located inside the BAG building polygons. This makes it possible to create links or spatial joins between the building polygons and residential objects and, as a result, to give information on

[6] LV is an abbreviation for "Landelijke Voorzieningen", which could be translated as "national services or facilities".

the type of building. The building polygons do themselves not contain information on the type or usage of these buildings.

Useful object classes	Useful attributes	Geometry type
Buildings	Geometry	Polygons
Residential objects	Purpose of use: - Residential function - Retail trade function - Gathering function - Accommodation function - Health function - Industry function - Sports function - Office function - Education function - Prison/Cells function - Other function	Point
	Status: - buildings process started - building permit granted	Point

Table 3.1: Useful object classes and attributes of the BAG dataset, combined with the geometry type (Ministerie van VROM 2009a, p. 75). See appendix B for the Dutch terms.

3.1.3 Mappings with culture codes

Figures 3.2 and 3.3 show conceptual mappings between the BAG dataset and the culture code dataset. Figure 3.2 shows that the BAG buildings can help to determine which cadastral parcels should have a built-up culture codes and which should have a non-built-up culture code. Figure 3.3 shows how the purposes of use of residential objects provide information for defining the right built-up culture codes, for those cadastral parcels with built-up elements located on them. There is a difference between direct and indirect mappings. The arrows with thick red lines represent 'direct' mappings. These are relationships that do not need additional information to create the culture codes. For example, for the culture code "Residential" only the purpose of use "residential" is needed. The 'indirect' mappings, represented by the arrows with dotted red lines, are relationships that need additional information in order to be able to create the culture code. An example of an indirect mapping is the mapping between the purpose of use "residential" and the culture code "Residential (apartment)". In order to be able to define that apartments are located on a cadastral parcel containing residential objects with the purpose of use "residential", additional spatial information on apartments is needed, as this is not included in the BAG dataset.

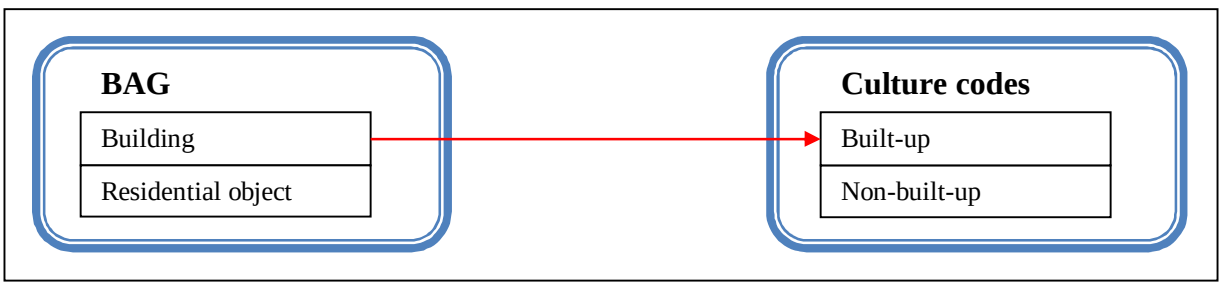


Figure 3.2: Mappings between BAG components and built-up or non-built-up code in the culture code dataset.

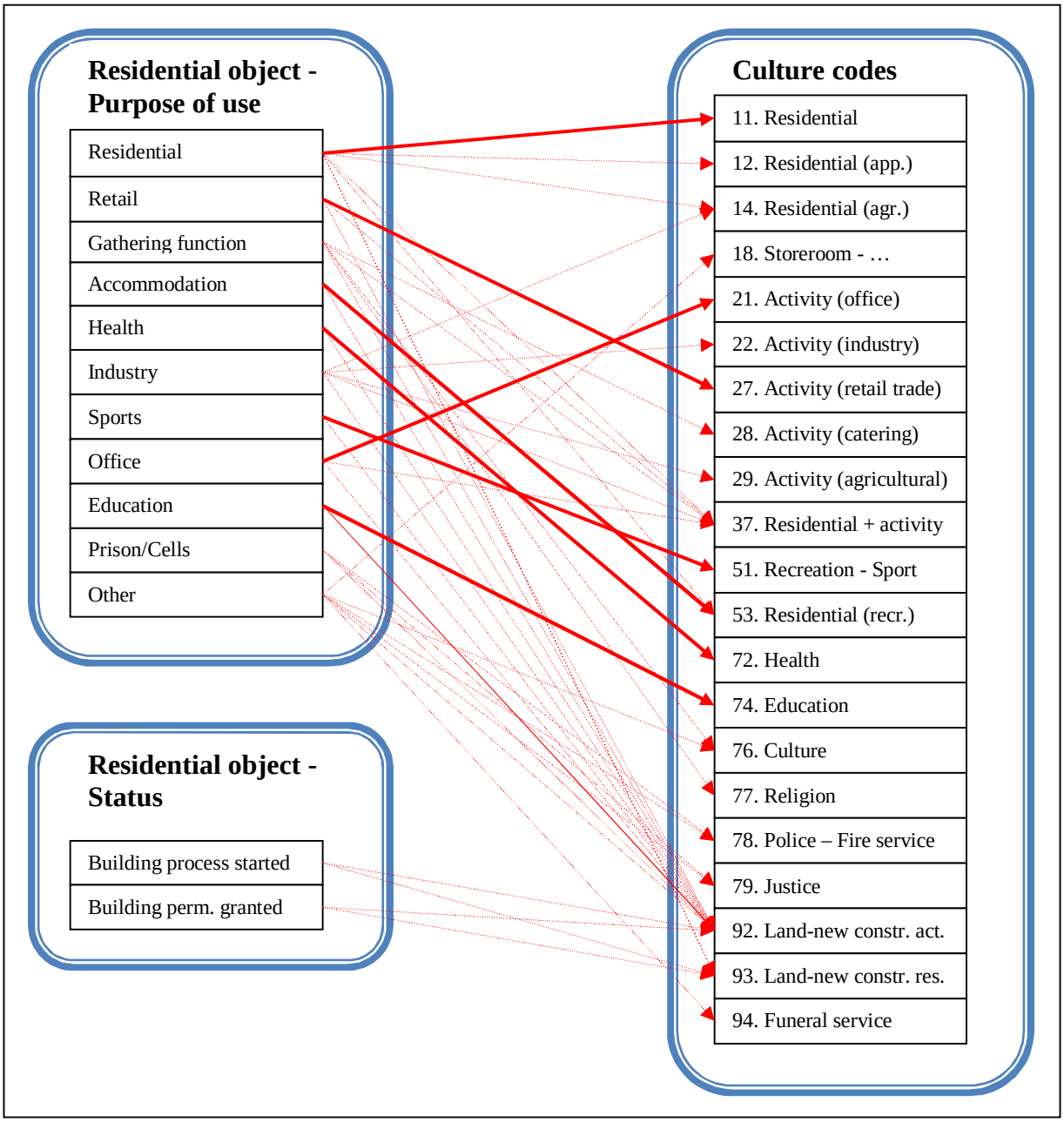


Figure 3.3: Mappings between purpose of use of BAG objects and culture codes.

3.1.4 Quality and up-to-dateness

With the BAG dataset being part of the Dutch system of key registers, a certain level of quality should be guaranteed. Audits are to be carried out regularly, in order to make sure this level of quality is indeed achieved by the Dutch municipalities. However, Kooij (2011) speaks out a number of doubts concerning the quality of the BAG dataset. These doubts include:

- The Dutch central office of statistics (CBS) has conducted an analysis on the quality and limitations of the BAG dataset. They concluded that the dataset has quite a few limitations concerning the content. However, they did not provide information on the type of limitations observed.
- The CBS has also stated that the data of the BAG dataset differ to a large extent between the municipalities.

- The Waarderingskamer (institution concerned with taxation of real estate objects) has indicated that the surfaces included in the BAG dataset differ largely from those included in the WOZ[7] dataset.
- The quality of the geometry attribute of the BAG dataset is relatively good, compared to the quality of other attributes. However, questions exist on the quality and completeness of the dataset concerning the underground BAG elements.
- Users of the BAG dataset are obliged to notify municipalities if they doubt the correctness of certain data. The municipalities do have to take steps after receiving such a notification. In some cases, the issue can be solved within a number of days. If more research is needed, the municipalities have six months to do so. During the research period, the status of the feature is set at "in research". This is a limitation to the up-to-dateness, moreover because it does not become clear from the dataset which attribute value is in research; the status is given to the whole feature
- As the BAG dataset is part of the Dutch system of key registers, there are strict rules on the up-to-dateness of the data, which is supervised by the central government. New building permissions, changes in geometry, as well as other changes in attributes of the dataset, have to be included within a number of days. However, as most municipalities have only started to send their data to the LV in 2010 and are now starting their maintenance phase, the deadlines are in many cases not (yet) met by the municipalities. In some cases municipalities have connected to the LV, but still have a lot of work to do before having their data up-to-date. They have "frozen" their dataset for a while and are now working on getting it up-to-date.

When using the spatial data quality criteria (explained in section 2.3) one gets a good overview of the general quality of a spatial dataset. For the BAG dataset, table 3.2 shows the score of the BAG datasets on the eleven spatial data quality criteria. The scores are based on the interview with the BAG specialist and the documentation on the dataset.

Spatial data quality criterium	Score	Explanation
Lineage	++	The BAG includes source documents for every change made to the dataset's features.
Positional accuracy	+	In general good (20cm for urban areas and 40cm for rural areas), but doubts on positional accuracy of underground BAG objects.
Attribute accuracy	-/+	Doubts exist on the quality of the attributes, also because of the variation in quality (see below)
Completeness	-	Many municipalities have work left to complete.
Logical consistency		[not known]
Semantic accuracy	+	Semantics are well-defined in the dataset's catalogue.
Temporal information/quality	++	The BAG contains information on the changes made to the dataset's features and contains historical data.
Usage, purpose, constraints		[not known]
Variation in quality	-	Large differences in quality between municipalities.
Meta-quality		[not known]
Resolution		[not known]

Table 3.2: Score of the BAG dataset on the eleven spatial data quality criteria (Kooij 2011; Ministerie van VROM 2009a, p. 45).

[7] The WOZ (abbreviation for Wet Waardering Onroerende Zaken) dataset contains information on real estate objects, including financial taxation, surface and building years.

3.1.5 Strengths and weaknesses

To sum up, table 3.3 presents the strengths and weaknesses of the BAG dataset. These strengths and weaknesses are used in the decision on which spatial datasets to use for the production of the culture codes dataset.

Strengths	Weaknesses
+ Dataset covers the whole surface of the Netherlands.	- Not completely completed yet by all municipalities.
+ No payment by Kadaster for using the dataset.	- Variation in quality between municipalities.
+ Up-to-date spatial data.	- Doubts on completeness looking at underground elements.
+ Audits should guarantee a basic level of quality.	- When having status "in research" it is not clear for which attribute value(s) this counts.
+ Accurate geometry for all object classes.	- Buildings can contain a large number of residential objects with sometimes different purposes of use.
+ Contains historical data.	

Table 3.3: Strengths and weaknesses of the BAG dataset.

3.2 – BRP: Dutch key register of agricultural parcels

3.2.1 Introduction

Unlike the BAG dataset, the BRP (visualized in figure 3.4) is no official key register of the Dutch government, although it is called a key register by the Dutch Ministry of Agriculture. The dataset contains information on agricultural parcels and nature areas. Examples of information covered by the dataset are the owner and the size of the parcels. More important in the context of this research, is the information on the use of the parcels. Although, in the case of the culture codes, it might be sufficient to store information on whether an agricultural parcel is used as grassland, crop-growing, greenhouses, etc., it would even be possible to make a distinction between different agricultural land use types by using the key register of agricultural parcels, which includes information on the crops grown on a certain parcel (Dijkstra 2005; Ministerie van LNV 2011).

3.2.2 Useful classes and geometry type

At the moment the Dutch Ministry of Finances, Agriculture and Innovation is in a transition phase when the Dutch key registers of parcels (BRP[8]) is concerned. Until now, the information included in the register was obtained from farmers who were obliged to give information on the use of their parcels every year. In the near future, the ministry wants to obtain the information from external – in their view, more reliable – resources, e.g. external source datasets. During the current transition phase, which lasts another six months from now, the ministry should decide on which external resources to use for the production of the BRP dataset and which method to apply. As long as the transition phase is not completed, the ministry is not willing to provide their data to third parties (Hendriks and Dijk 2011).

As a result, this chapter does not give a detailed description of the characteristics of the BRP dataset. Hendriks and Dijk (2011) have suggested to make use of the LGN dataset, which is described in section 3.5 of this document.

[8] BRP is an abbreviation for "Basisregistratie Percelen", which can be translated as "key registers of parcels".

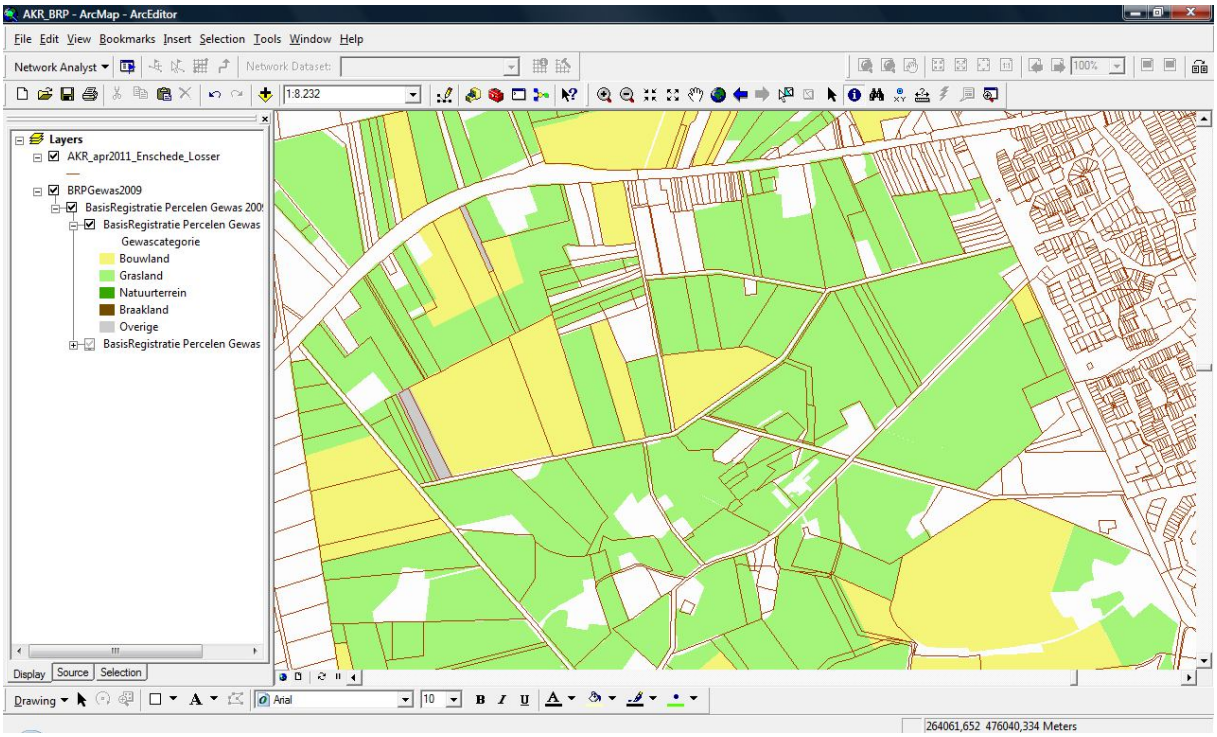


Figure 3.4: Extract from the web service of the BRP dataset imported in ArcGIS (GIS Service Agro.nl 2011).

3.2.3 Mappings with culture codes

Figure 3.5 shows how the BRP dataset could help to create a number of culture codes. On the left side the figure shows the useful BRP classes. On the right side the different culture codes that could be created with these BRP classes are shown. The arrows between the BRP classes and the culture codes show the relations between them.

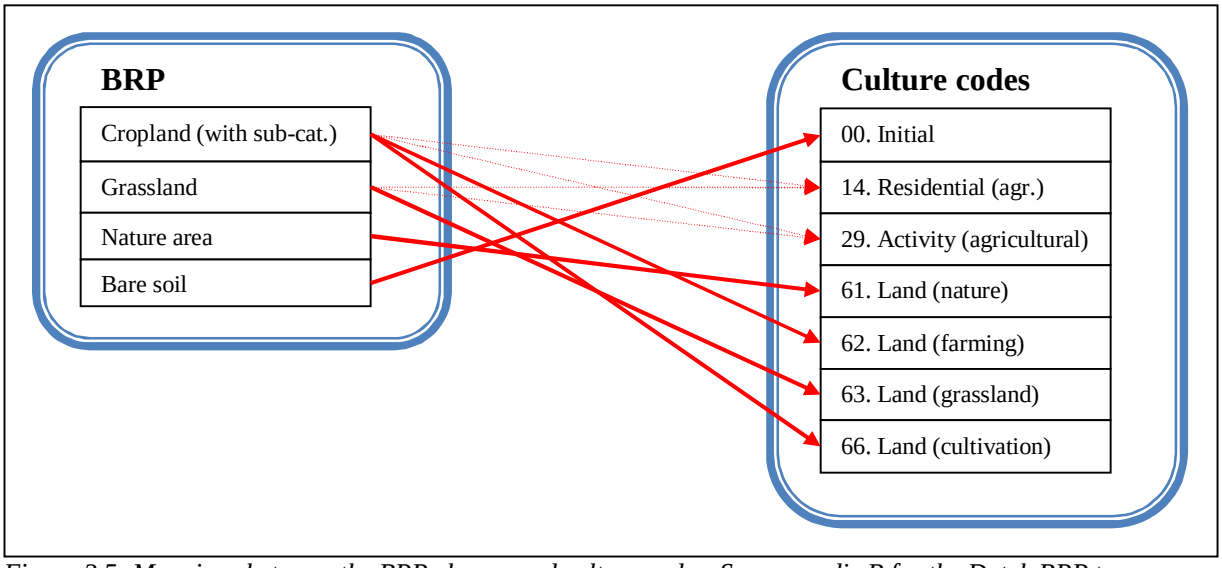


Figure 3.5: Mappings between the BRP classes and culture codes. See appendix B for the Dutch BRP terms.

3.2.4 Quality and up-to-dateness

Because of the transition phase the dataset is not available at the moment and the Ministry of Agriculture is not willing to give detailed information on the quality of the dataset. Therefore, no detailed quality description of the dataset is included in this thesis. However, the fact that the Ministry of Agriculture is looking for a more reliable, new production method of the dataset, suggests that the quality is not satisfying at the moment.

3.2.5 Strengths and weaknesses

Not much can be said about the strengths and weaknesses of the BRP dataset. At the moment, the most important weakness is that the dataset is not available. A strength of the BRP dataset could be its up-to-dateness, when compared to the TOP10NL and LGN (see sections 3.3 and 3.5), as the use of the agricultural parcels is updated every year. Another strength could be the large amount of details on use of the agricultural parcels.

3.3 – TOP10NL/BRT

3.3.1 Introduction

The BRT is the Dutch key register of topography. It contains a number of topographic datasets with a smaller scale than the scale the GBKN dataset is used for (1:500 to 1:5.000, see section 3.4). The TOP10NL – visualized in figure 3.6 and with its name referring to the map scale of 1:10.000 – is one of the topographic datasets being part of the key register. The dataset is produced by Kadaster (in Zwolle). The main source for the information needed is aerial photography. The aerial photographs are interpreted by cartographers working for Kadaster. Next to aerial photography, panoramic pictures (cyclorama pictures) are used for gathering the needed information. Such panoramic pictures are taken from a driving car and show the situation from the ground-level instead from the air, in case of the aerial photographs (te Winkel 2011b).

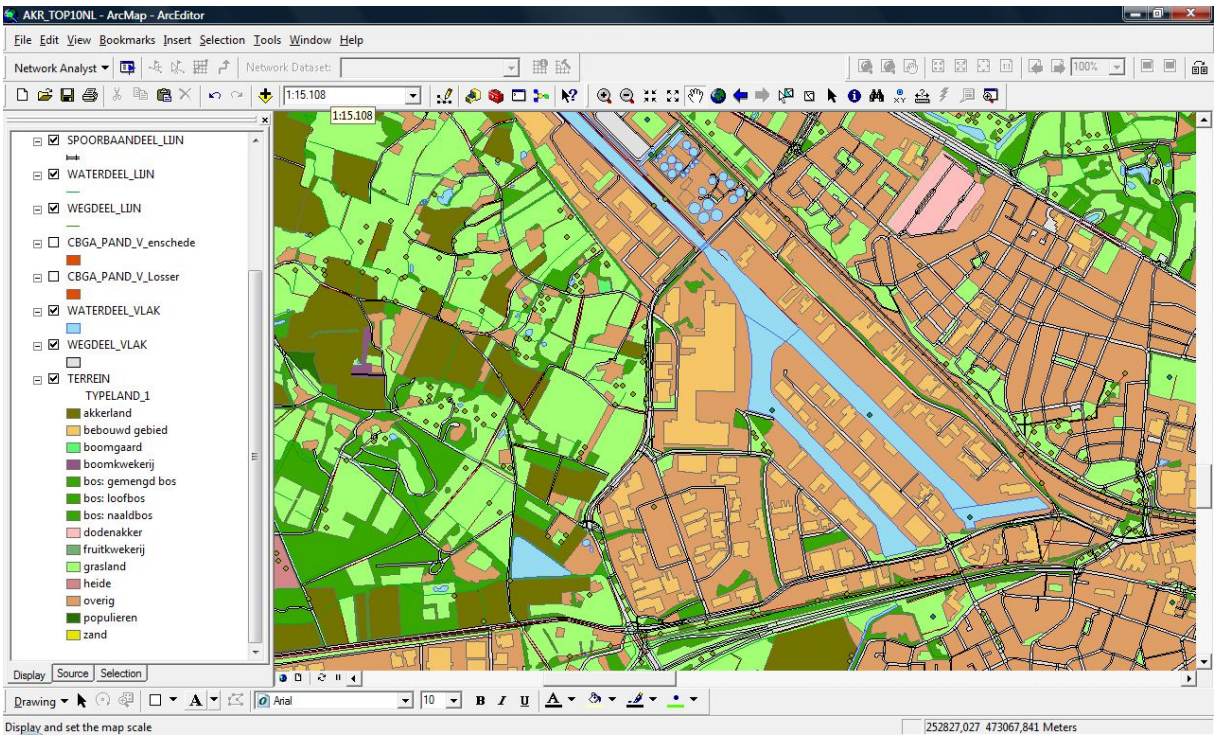


Figure 3.6: Visualization of the TOP10NL dataset in ArcGIS.

Figure 3.7 shows a simple UML diagram for the NEN3610 standard. This is a Dutch standard for geographical objects (NEN 2011). Many of these geographical object types are included in the TOP10NL dataset: terrain, water, roads, railroads, buildings, cultural landscape objects, administrative areas, functional areas and geographical areas. These nine object classes of the TOP10NL offer useful information for the production of the culture codes dataset. Section

3.3.2 offers more information on the usefulness and geometry types distinguished between in the TOP10NL dataset.

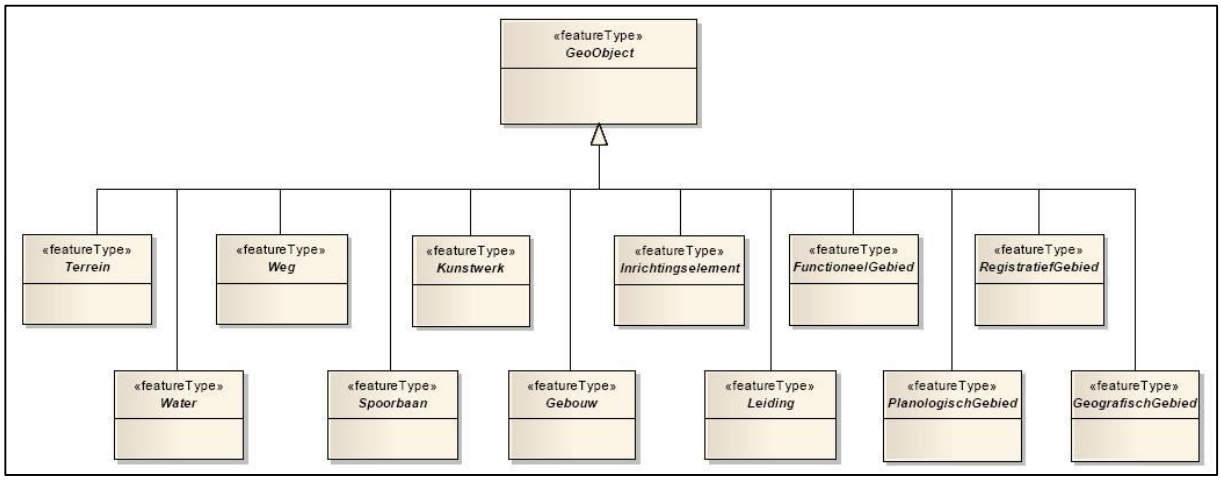


Figure 3.7: NEN3610 UML diagram, containing most of the TOP10NL object classes (NEN 2011, p. 35).

3.3.2 Useful classes and geometry type

Table 3.4 presents the useful TOP10NL classes, attributes and attribute values for the production of the culture codes dataset. Some attributes which could be used for the production of the culture codes are not obligatory, as the dataset distinguishes between obligatory and optional attributes (Bakker, Bruns, and Storm 2005, p. 14; te Winkel 2011b). This means that not in all cases the information is available in the dataset, seriously limiting the completeness of these attributes (te Winkel 2011b). Table 3.4, therefore, also includes a column indicating whether attributes are obligatory or optional.

The TOP10NL dataset contains three types of geometry: polygons, lines, and points. Depending on the object class and attribute, one or more of these geometry types is/are possible (Bakker, Bruns, and Storm 2005, p. 15; te Winkel 2011b). The list of interesting objects, attributes and attribute values in table 3.4 also provides information on the geometry, as this has a strong influence on the way it could be used for the purpose of the research described in this document.

It is also important to mention that the dataset contains both “partitioning” and “overlapping” objects. Partitioning objects are water and road sections, as well as terrain objects. Together these objects are covering the whole surface of The Netherlands, ensuring there are no “white” areas, which are not covered by one of these objects (Bakker, Bruns, and Storm 2005, p. 16; te Winkel 2011b). The other objects are “overlapping” objects, meaning they are placed mainly on top of terrain objects, and sometimes on top of road or water objects. Also road, water and terrain objects can be overlapping (Bakker, Bruns, and Storm 2005, pp. 16-17). For example, a road section could overlap a water object, as a bridge could be built to cross the river or lake.

Finally, the features stored in the TOP10NL dataset do not geometrically coincide with the cadastral parcels. Features are defined by their land use and do not take into account cadastral parcel borders. Bordering areas having the same land use type are aggregated to one feature, unless there is a clear division visible “in the field” (te Winkel 2011b).

Object class	Attribute *Attribute value*	OPT/ OBL	Point geometry	Line geometry	Polygon geometry
Road section	Main use *Bus traffic* *Horse riding* *Air traffic* *Parking* *Park + ride* *Carpool site* *Mixed traffic* *Fast traffic*	OBL	x	x	x
Railway section	Type of railway *Train* *Tram* *Metro/Subway* *Mixed*	OBL	x	x	
Water object	Type of water *Water Stream* *Lake, fen, etc.* *Ditch* *Sea* *Tidal* *Well*	OBL	x	x	x
	Physical appearance *In a sluice*	OPT		x	x
	Function *Port* *Natural pool* *Fish ladder* *Water treatment* *Swimming pool*	OBL	x	x	x
Building	Type of building *Visitor centre* *Crematory* *Dock* *Powerplant* *Plant/Factory* *Pumping station* *Prison* *Hotel* *Housing block* *Chapel* *Church* *Nuclear powerplant* *Monastery* *Clinic* *Artificial ice track* *Horse riding school* *Metro station* *Military building* *Motel* *Museum* *Parking garage* *Police office* *Petrol station* *Psychiatric clinic* *Recreation centre* *Lifeboat building* *Religious building*	OPT			x

	Coach-house *School* *Sports building* *Stadium* *Train station* *University* *Traffic tower* *Road restaurant* *Shipyard* *Hospital* *Swimming pool*				
Terrain	Land use type *Landing-stage/Pier* *Cropland* *Tree nursery* *Forest* *Fruit farm* *Grassland* *Heath/Moor*	OBL			x
"Cultural landscape object"	Type of object *Landing-stage/Pier* *Gas extraction* *Memorial site/Monument* *Pumping station* *Helicopter platform* *Power line* *"Hunebed"* *Chapel* *Shooting range* *Sluice door* *Water barrier* *Station* *Dam*	OBL	x	x	
Administrative area	Type of administrative area *National park*	OBL	x		x
Geographical area	Type of geographical area *Forest area* *Dune area* *Moor/Heath area* *Lake, fen, etc.* *Mudflat* *Sea* *Sea arm*	OBL	x		x
Functional area	Type of functional area *Arboretum* *Business area* *Cemetery* *Bungalow park* *Camping place* *Caravan park* *Circuit* *Crossing track* *Zoo* *Golf course* *Burial mound* *Port* *Helicopter landing site* *Yacht harbor*	OBL	x		x

	Karts track *Military area* *Mine* *Mussel bed* *Nature area* *Open air museum* *Open air theatre* *Park* *Recreational area* *Hippodrome* *Ski slope* *Sluice* *Sports site* *Dump* *Tennis course* *Garden centre* *Airport* *Allotment* *Shipyard* *Ice track* *Water treatment*

Table 3.4: Object classes of the TOP10NL dataset with attributes, attribute values and geometry type (OPT = optional, OBL = obligatory) (Bakker, Bruns, and Storm 2005, pp. 23-48; Kadaster 2007, 2011b).

3.3.3 Mappings with culture codes

As illustrated by table 3.4, a long list of attribute classes defined in the TOP10NL dataset offer interesting information for the production of the culture code dataset. As a result, a large number of mappings can be visualized for the TOP10NL dataset (figures 3.8 through 3.15). In case the TOP10NL offers useful data for a culture code that does no longer exist, this culture code is presented in red.

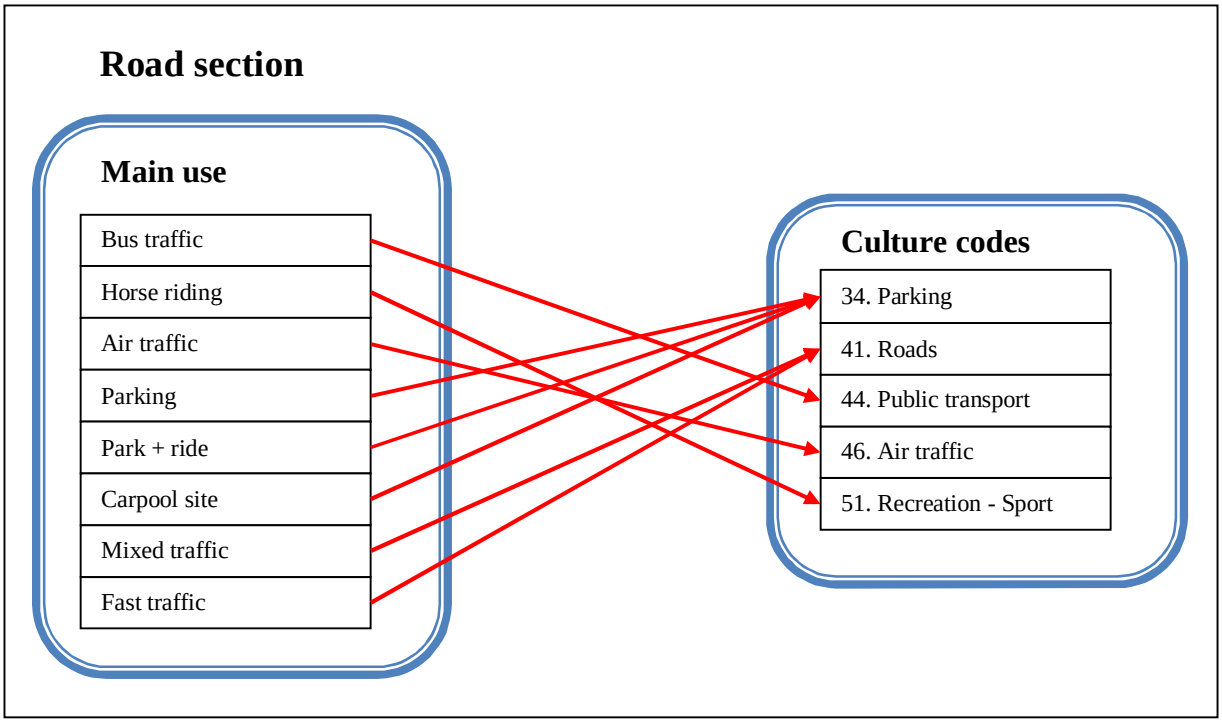


Figure 3.8: Mappings between road sections (TOP10NL) and culture codes.

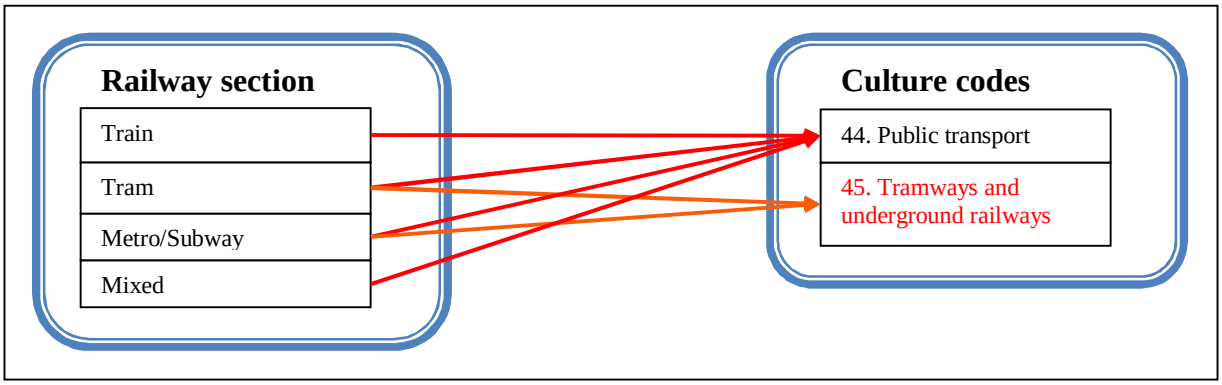


Figure 3.9: Mappings between railway sections (TOP10NL) and culture codes.

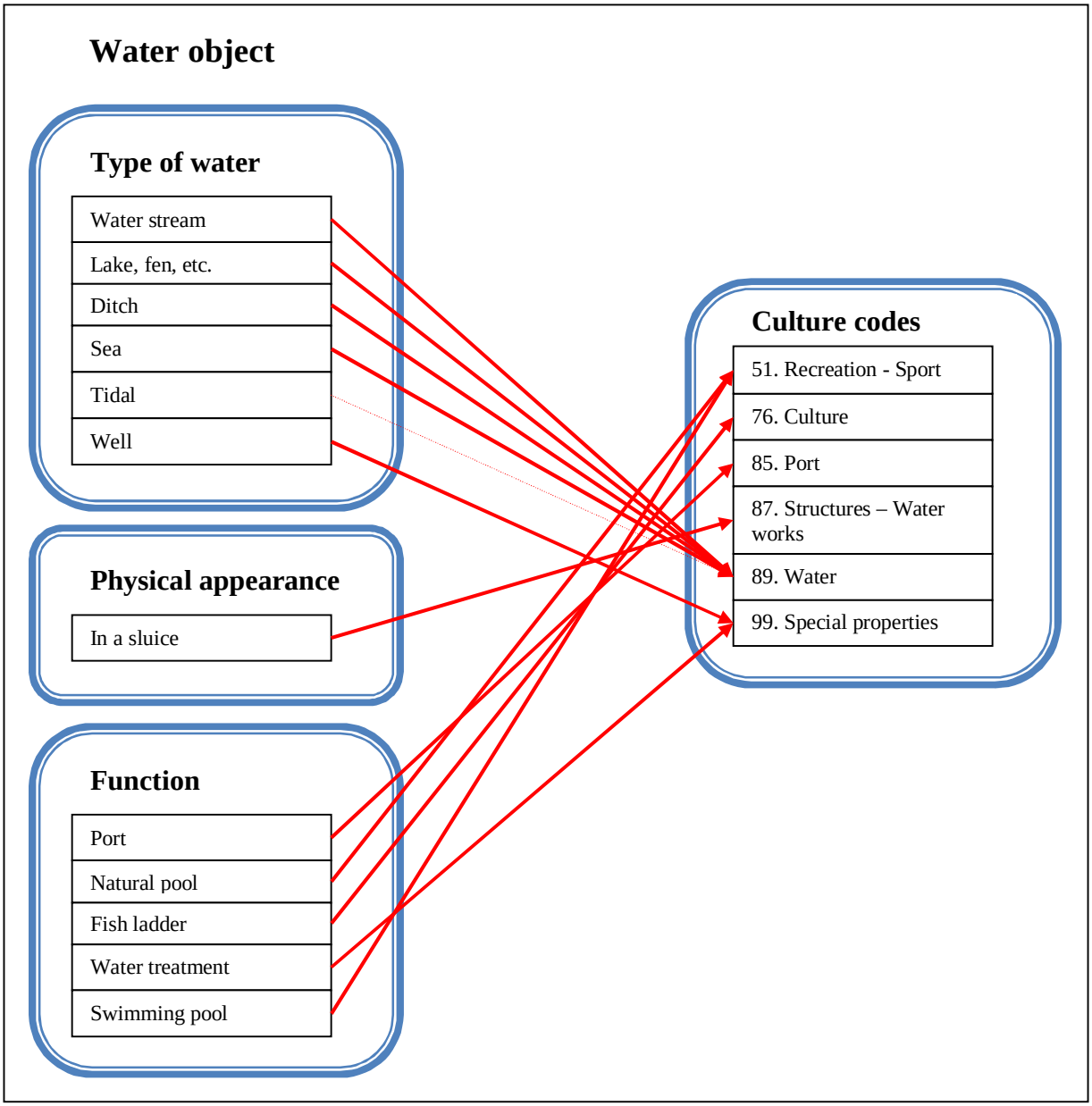


Figure 3.10: Mappings between water objects (TOP10NL) and culture codes.

Figure 3.11: Mappings between building types (TOP10NL) and culture codes.

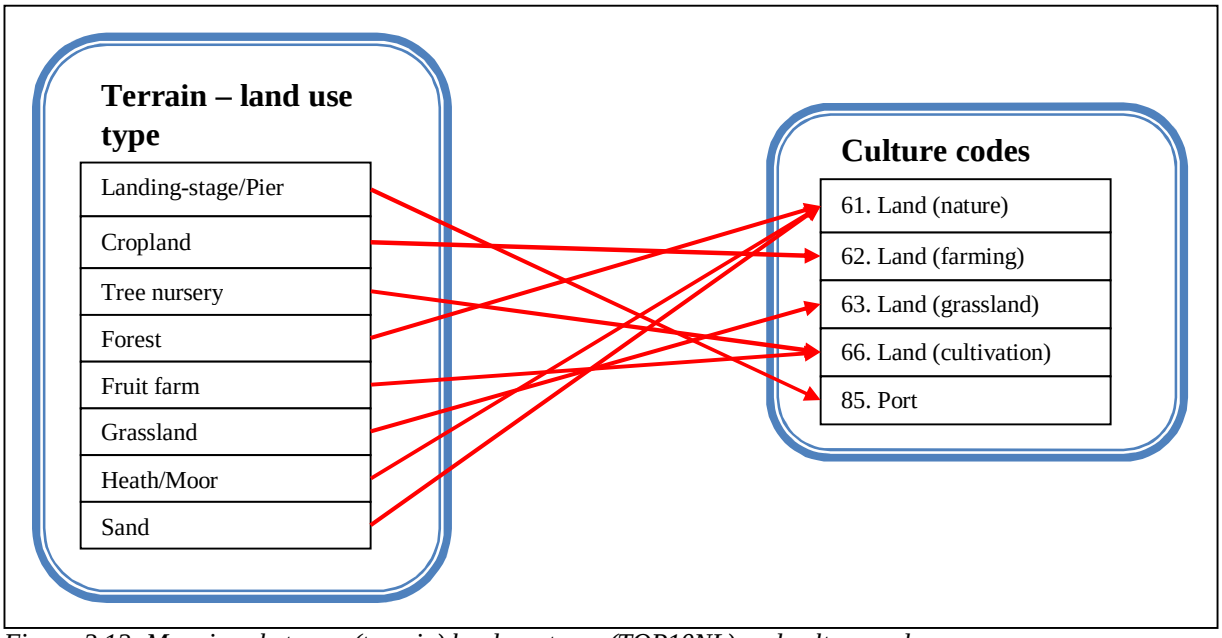


Figure 3.12: Mappings between (terrain) land use types (TOP10NL) and culture codes.

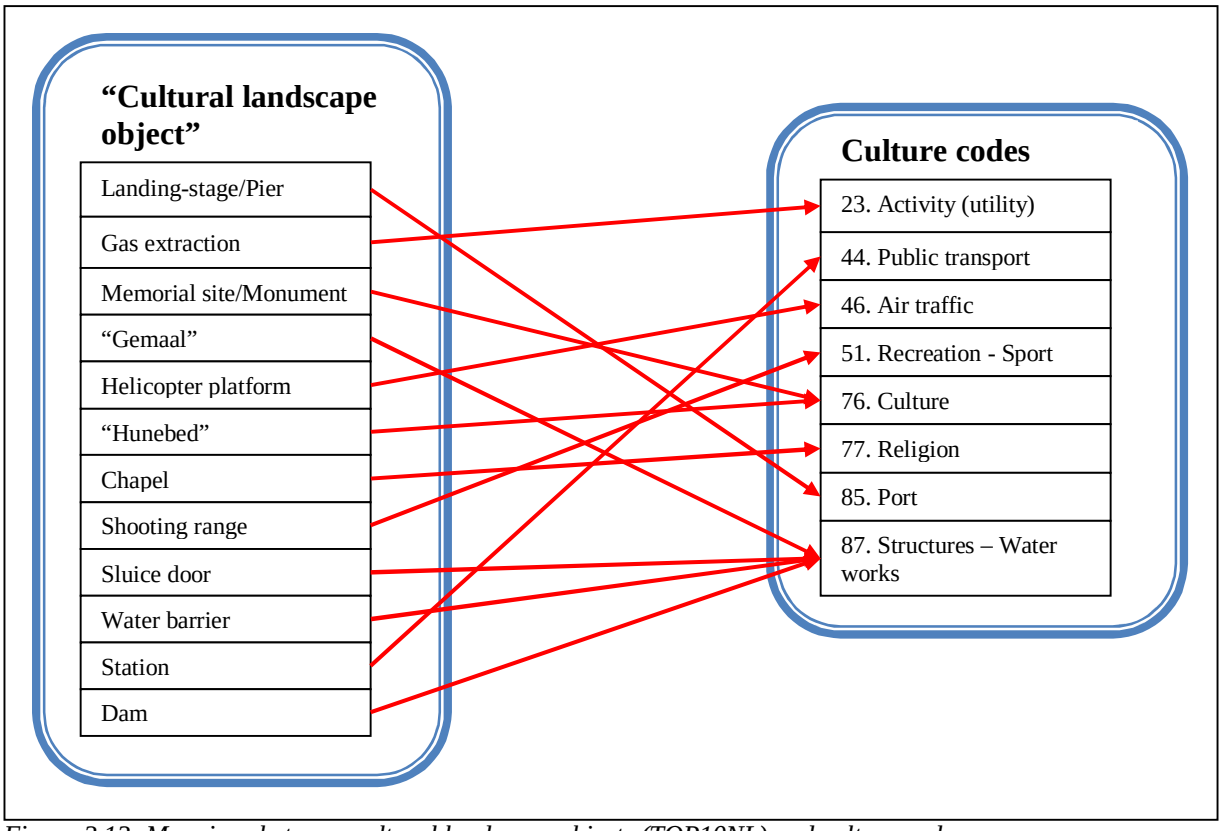


Figure 3.13: Mappings between cultural landscape objects (TOP10NL) and culture codes.

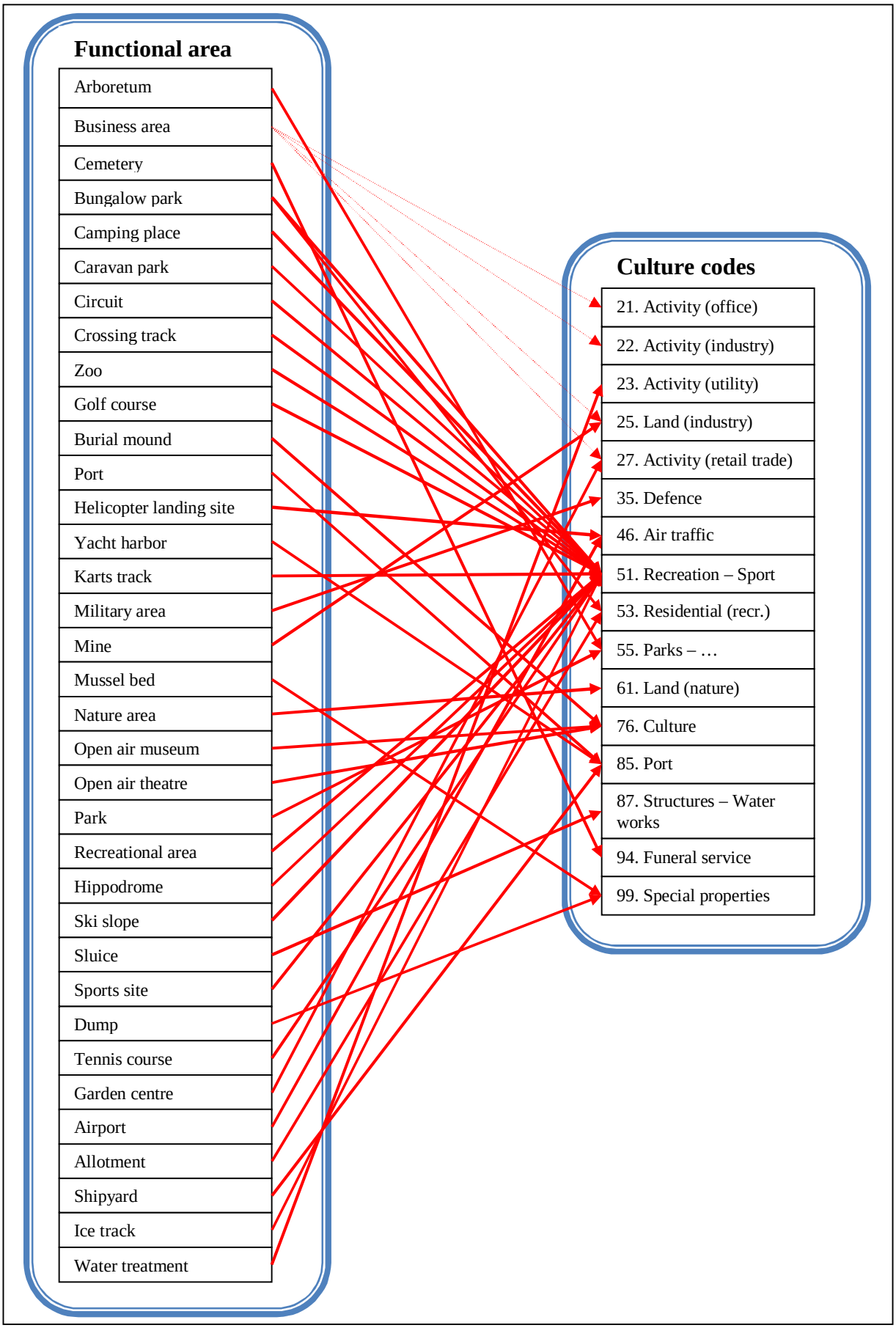


Figure 3.14: Mappings between functional areas (TOP10NL) and culture codes.

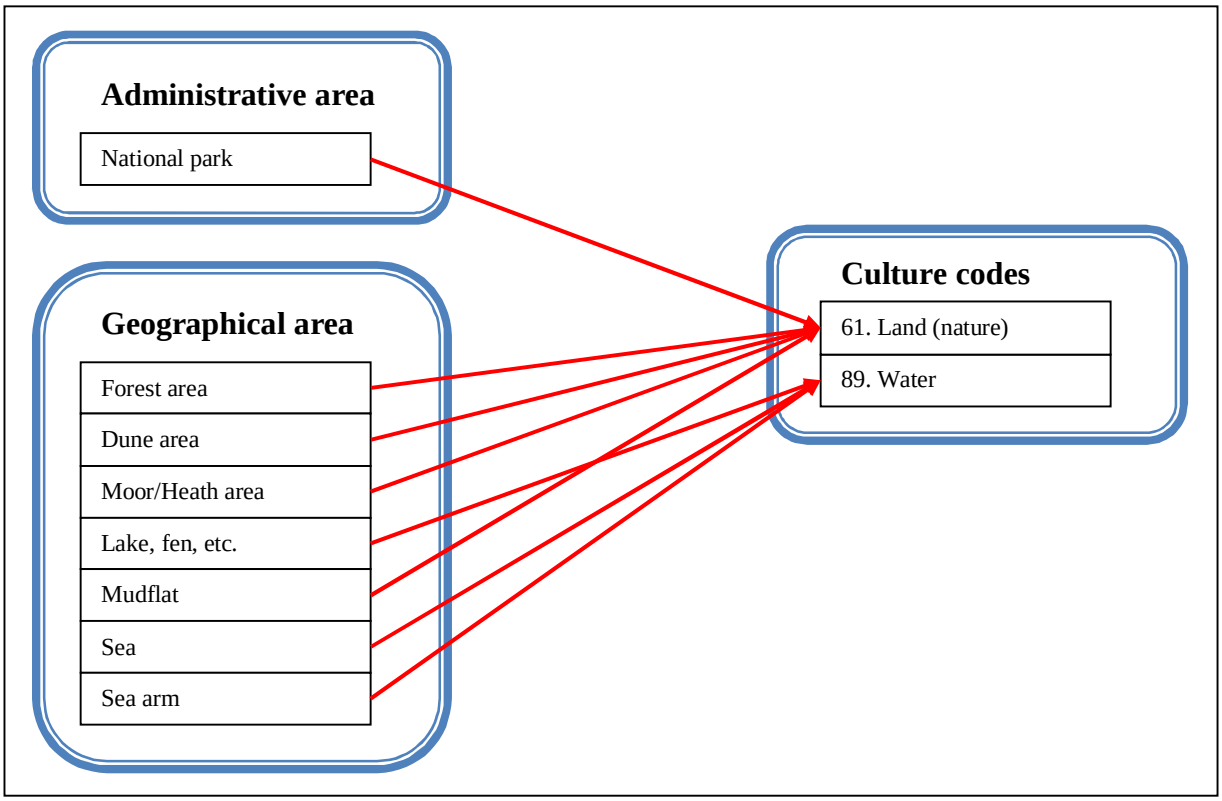


Figure 3.15: Mappings between administrative and geographical areas (TOP10NL) and culture codes.

3.3.4 Quality and up-to-dateness

The TOP10NL is updated every two years, meaning that every year one half of The Netherlands is updated and that data cannot be more than two years old (te Winkel 2011b). For example, if the western part of The Netherlands is updated in 2011, the next update round would be in 2013. The eastern part would then be updated in 2012 and 2014.

Unfortunately, no numbers exist on the quality of the dataset (te Winkel 2011b). However, key registers do need to have a certain level of quality and audits, carried out by the central government, should guarantee this level of quality. Furthermore, Kadaster does not receive many complaints about the content of their dataset from the users. This could mean the dataset is qualitatively good, but could also mean that the users are not reporting mistakes in the data (te Winkel 2011b). An indication of the quality of the TOP10NL dataset is presented by means of table 3.5, which includes the score of the dataset on the spatial data quality criteria.

Spatial data quality criterium	Score	Explanation
Lineage	-/+	Not really available, but TOP10NL is updated every two years.
Positional accuracy	-/+	Expected to be around three meters.
Attribute accuracy	+	Like completeness, should be around 95%.
Completeness	+	95% complete.
Logical consistency	+	Taken care of by formulating and applying validation rules.
Semantic accuracy	++	Detailed description of the semantics per (attribute) class.
Temporal information/quality	+	Information not older than two years. Historical data are stored in core database, but not (yet) made available for end-users.

Usage, purpose, constraints	+	Detailed description of usage. Used for wide range of purposes.
Variation in quality	-	Variation is possible, because of the amount of different people working on the product. However, by making sure people are not always assigned to the same area, the product quality becomes more equal.
Meta-quality	+	Internal and external quality validation.
Resolution	-	A few meters. 10 cm resolution aerial photographs are used, but changes are only made to the dataset when difference between photograph and TOP10NL are more than 5 meters.

Table 3.5: Score of the TOP10NL dataset on the eleven spatial data quality criteria (te Winkel 2011a).

3.3.5 Strengths and weaknesses

From the previous sub-sections it can be concluded that the TOP10NL seems to be a useful spatial dataset for the production of the culture code dataset. Table 3.6 contains the most important strengths and weaknesses of the dataset.

Strengths	Weaknesses
+ Dataset covers the whole surface of the Netherlands	- Updated only every two years.
+ No payment by Kadaster for using dataset.	- Some attributes are optional.
+ High level of detail in attribute values.	- Overlapping object classes and attributes are possible.
+ Basic level of quality should be guaranteed as the dataset is part of a Dutch key register.	

Table 3.6: Strengths and weaknesses of the TOP10NL dataset.

3.4 – GBKN/BGT

3.4.1 Introduction

In the previous section, information has been provided on the TOP10NL dataset, which is a small-scale topographical dataset. In this section another topographic dataset is presented: the large-scale base map of the Netherlands, the GBKN. In the future, the GBKN dataset is to be transformed into the BGT, the key register of large-scale topography, which will be part of the Dutch system of key registers. However, it is assumed that it will take until 2014 or 2015 before this new key register will be completed (Peersmann 2009; van Rossem 2009). The GBKN (visualized in figure 3.16) is described as follows:

- A digital topographic map with a defined minimum content and precision, which contains the most important topographic elements visible “in the field” (buildings, roads, water objects, etc.);
- A large-scale map: the map is used at geographic scales between 1:500 and 1:5000. Because of this large scale, many details can be displayed in the map;
- A base dataset, serving as a background map and used for many divergent applications. Users can, if necessary, add additional information to the map;
- Available for every part of the Netherlands.

(LSV GBKN 2009)

Figure 3.17 shows the object classes distinguished between in the BGT.

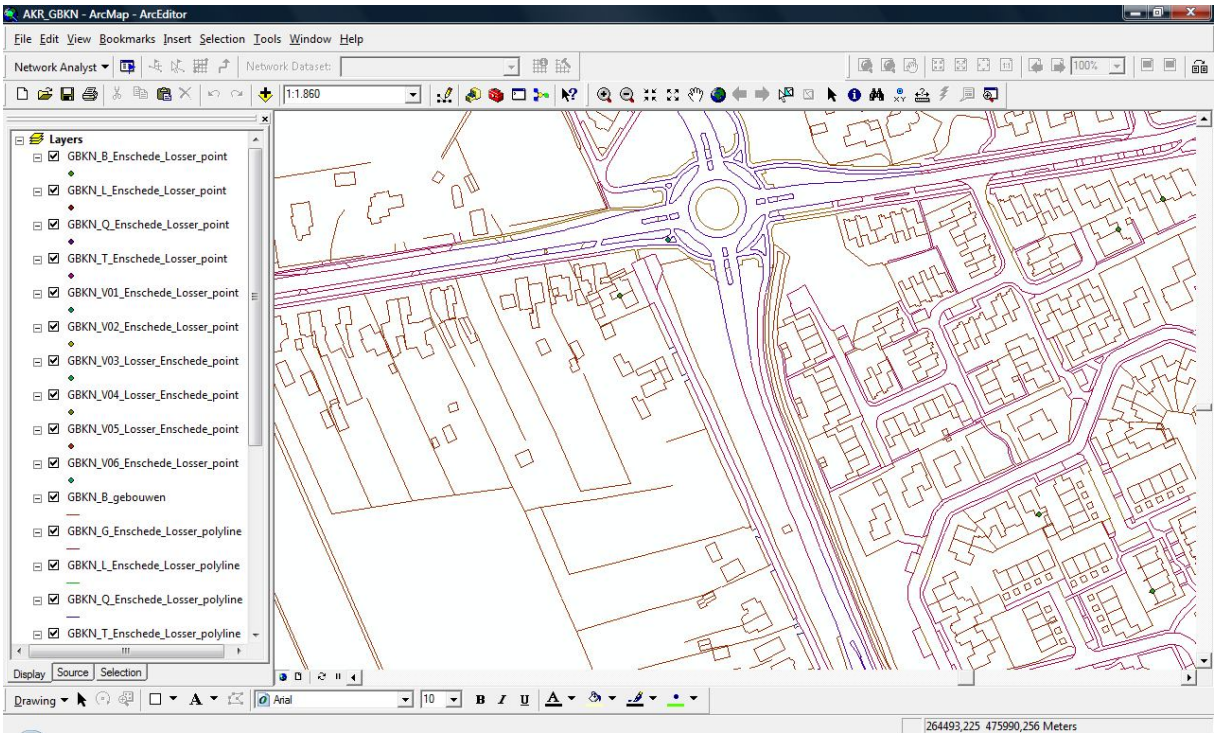


Figure 3.16: Visualization of the GBKN dataset in ArcGIS.

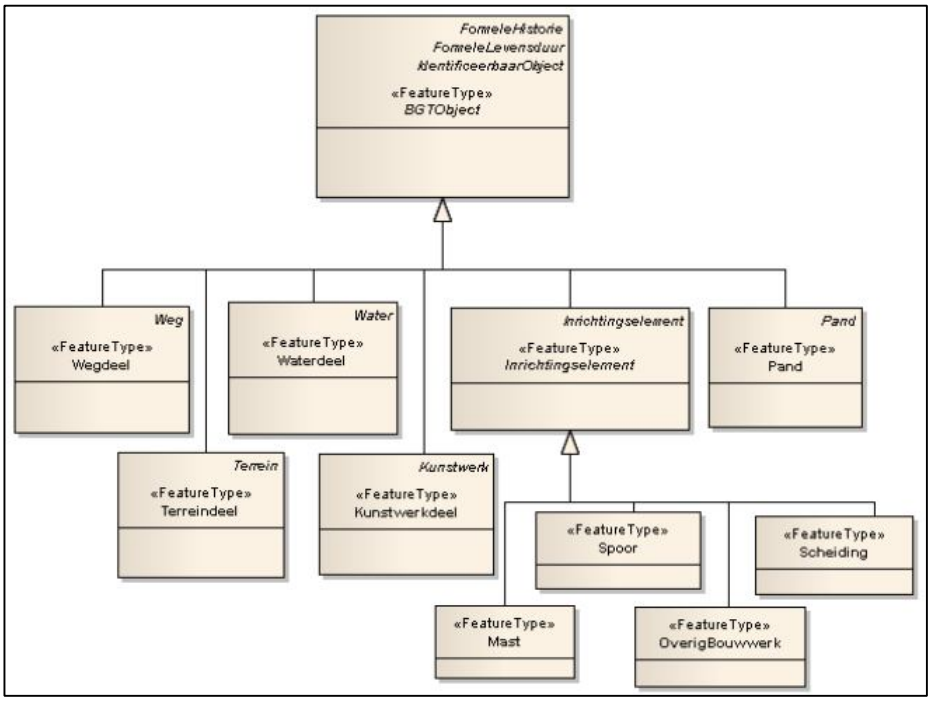


Figure 3.17: UML class diagram of the BGT object classes (Programma BGT 2010, p. 8).

The GBKN dataset is produced by means of land surveying techniques. A number of producers of the GBKN/BGT dataset exist, among others: municipalities, the rail company and the water boards. However, they have organized the production of the dataset differently; some produce the data themselves, others pay commercial companies to measure and position the objects to be included in the dataset. In the GBKN this diversity in ways of production has lead to a range of different “flavours” concerning the semantics and, as a result, the contents of the dataset and map. The introduction of the BGT should help to decrease the level of

diversity and to create a “smooth” map, by clearly defining the what and how of the new dataset (van Rossem 2009).

3.4.2 Useful classes and geometry type

Although, when compared to the TOP10NL dataset, the GBKN and BGT have much more detail concerning the shape of objects due to the larger scale of the map, at attribute level they contain less details (Maessen 2011b). The BGT will contain a number of attributes per object class, which could be valuable for creating the culture codes. These attributes are listed in table 3.7, together with the geometry type of the object classes.

Like the TOP10NL dataset, the large-scale topographic dataset objects do not coincide with the cadastral parcels (Maessen 2011b). However, whereas in the TOP10NL dataset the objects are often larger than the parcels, in the GBKN and BGT it happens more often that the objects have a smaller size than the parcels. This can be explained by the fact that the GBKN and BGT show more details, as they are produced for a larger geographical scale.

Most important to mention in this section is the fact that the geometry of the GBKN differs largely from that of the BGT. Whereas the GBKN consists of line elements showing the borders of a certain object, the BGT consists of polygons representing these objects. Maessen (2011b) argues that this is one reason why the current GBKN dataset is not very useful for the production of the culture codes. The line elements make it difficult to calculate the percentage of a parcel containing, for example, a road section. Such a percentage might be necessary to apply decision rules when more than one land use type is found on a certain cadastral parcel. When the BGT dataset would be available, its usefulness would be much better (Maessen 2011b).

Object class	**Attribute** ***Domain value***	**Geometry type**
Road section	Type of road *Starting lane for air traffic* *Highway* *Main road* *Regional road* *Local road* *Street* *Path*	GBKN: line BGT: polygon
Water object	Type of water object *Sea* *Water stream* *Water surface*	GBKN: line BGT: polygon
Engineering object	Type of engineering object *Bridge* *Pumping station* *Platform (public transport)* *Sluice* *Dam* *Tunnel* *Wind turbine* *Weir* *Stairs*	GBKN: line BGT: polygon
Separation element	Type of separation element *Wall* *Quay/Pier wall* *Noise screen* *Dam wall* *Shore protection* *Hedge*	GBKN: line BGT: polygon

	Fence	
Other built elements	Type of built element *Open stall* *Settling tank* *Low transformer* *Basin* *Silo* *Landing-stage/Pier*	GBKN: line BGT: polygon

Table 3.7: Object classes of the GBKN/BGT dataset with attributes and domain values, which might be useful for production of the culture codes (Programma BGT 2010, pp. 32-64).

3.4.3 Mappings with culture codes

Figures 3.18 through 3.22 visually present the mappings between the GBKN classes and the culture code dataset. Per class distinguished in table 3.7 the relationships between sub-classes and individual cultural codes are drawn as arrows.

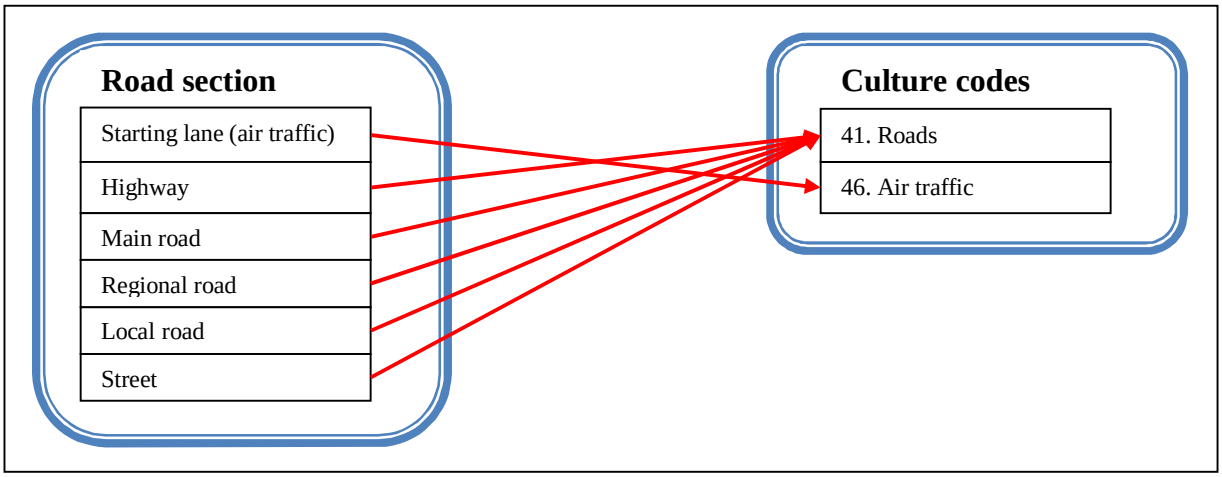


Figure 3.18: Mappings between road sections (GBKN/BGT) and culture codes.

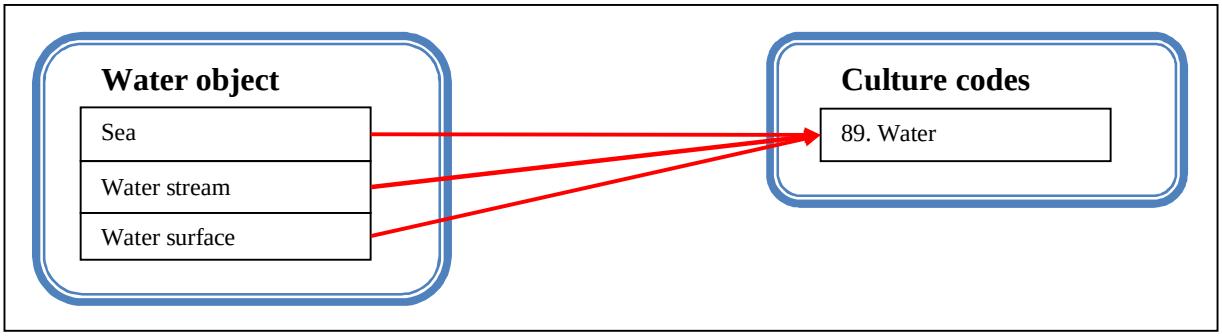


Figure 3.19: Mappings between water objects (GBKN/BGT) and culture codes.

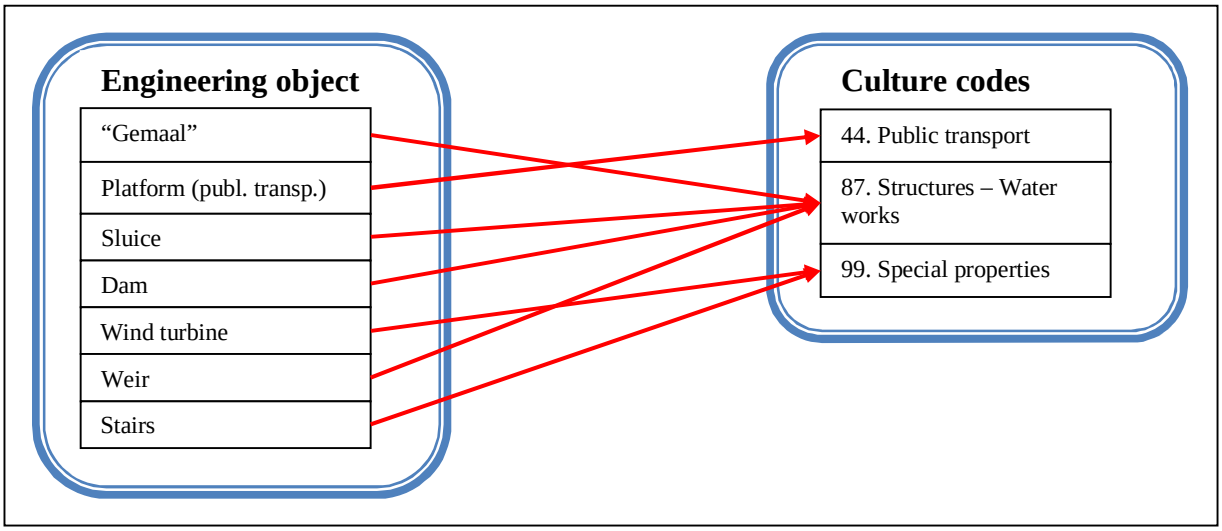


Figure 3.20: Mappings between engineering objects (GBKN/BGT) and culture codes.

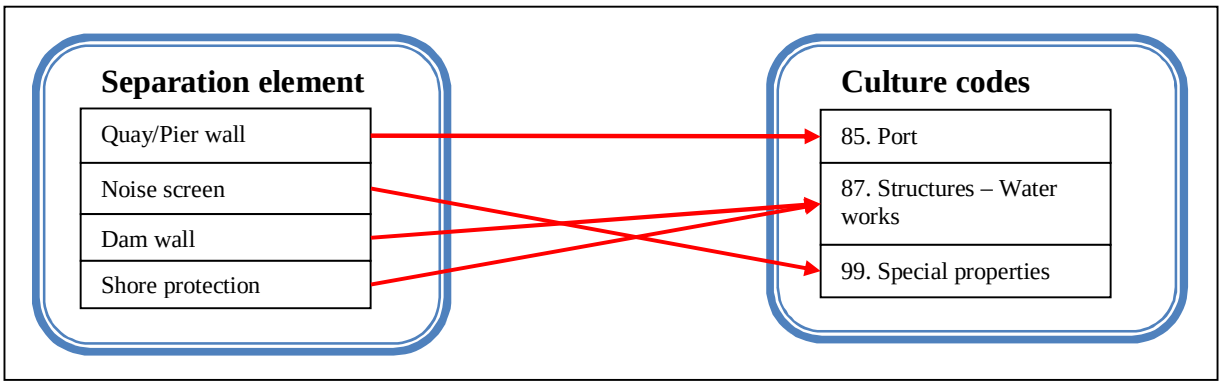


Figure 3.21: Mappings between separation elements (GBKN/BGT) and culture codes.

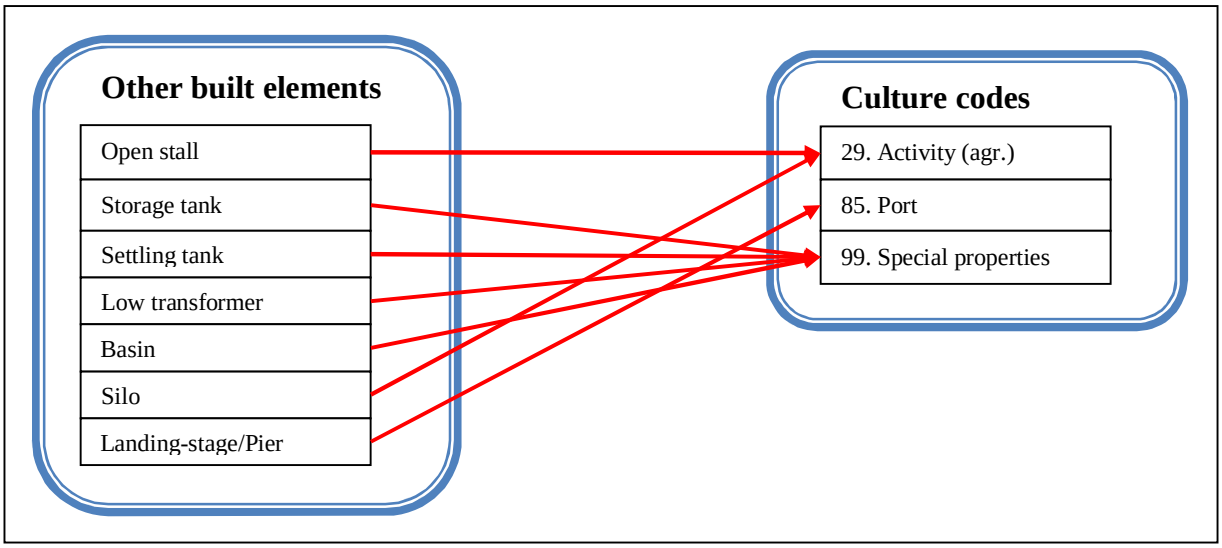


Figure 3.22: Mappings between other built elements (GBKN/BGT) and culture codes.

3.4.4 Quality and up-to-dateness

All the object classes of GBKN included in the mapping figures (see above) have an up-to-dateness of a maximum of 12 months. This means changes in objects have to be included in the dataset within 12 months. With the introduction of the BGT, for roads this will become 6 months (Programma BGT 2010, p. 20). This means, the up-to-dateness will increase a bit with the introduction of the BGT.

Maessen (2011b) seriously questions the quality of the current GBKN dataset. He argues the quality of the dataset is satisfying for places which are directly visible from the public roads, especially in urban areas. Outside cities and villages the GBKN is far from complete. With the creation of the BGT key register, the quality of the large-scale topography dataset should increase significantly. Table 3.8 gives the scores of the GBKN dataset on the spatial data quality criteria.

Spatial data quality criterium	Score	Explanation
Lineage	-	Not sufficiently taken care of.
Positional accuracy	++	Good, accuracy of 30 till 60 (depending on the type of object class).
Attribute accuracy	+	Quite good, but some complaints about attribute accuracy for buildings.
Completeness	-	Quite complete for objects directly visible from streets. Far from complete for areas outside cities and villages.
Logical consistency		[unknown]
Semantic accuracy	--	Not well taken care of. For many regulations/rules it is hard to be controlled whether they are applied.
Temporal information/quality	-	New information has to be added within 6 – 12 months (depending on which type of information).
Usage, purpose, constraints	+	Usage and purpose are the same; used by all governmental institutions. Financial constraints for non-governmental users.
Variation in quality	--	Large variation in quality. High quality in cities and villages; much lower quality and less complete in other areas.
Meta-quality	+	Ok, available in the dataset.
Resolution	++	High resolution, depending on the object class.

Table 3.8: Score of the GBKN dataset on the eleven spatial data quality criteria (Maessen 2011a).

3.4.5 Strengths and weaknesses

Table 3.9 presents the strengths and weaknesses of the GBKN dataset, as this is still the available dataset. With the introduction of the BGT the up-to-dateness for road sections will improve and the datasets will consist of polygons instead of line elements. The latter change will significantly increase the usefulness of the large-scale topographic dataset for the production of the culture code dataset.

Strengths	**Weaknesses**
+ Dataset covers the whole surface of the Netherlands.	- Unsatisfying quality outside cities and villages.
+ No payment by Kadaster for using dataset.	- Dataset consists of line elements.
+ High level of geometrical detail.	- Less details at attribute level than TOP10NL dataset.
+ Up-to-date (changes included within 12 months)	- Possible to have more than one object (class)/attribute at one parcel.

Table 3.9: Strengths and weaknesses of the GBKN dataset.

3.5 – LGN

3.5.1 Introduction

The LGN (Landelijk Grondgebruik Nederland – rural land use The Netherlands) dataset is produced by Alterra, a research institute related to the Wageningen University in the Netherlands. The LGN dataset (visualized in figure 3.23) covers the whole surface of the Netherlands and gives the user information on the land use on a certain location. As the main focus of the dataset is on the rural areas, the information for urban areas is very limited. However, the dataset does contain a wide range of interesting object classes defining the rural land use of certain locations.

The main information sources for the production of the dataset are satellite images and aerial photography. Remote sensing techniques have a significant role in the automated derivation of the land use types (Alterra 2011). Furthermore, external datasets (such as a map containing nature areas) are used for additional information on the land use.

The LGN dataset has to be obtained from Alterra. The payment fee is more or less €63.000 for the whole surface of the Netherlands, not including taxes and delivery costs.

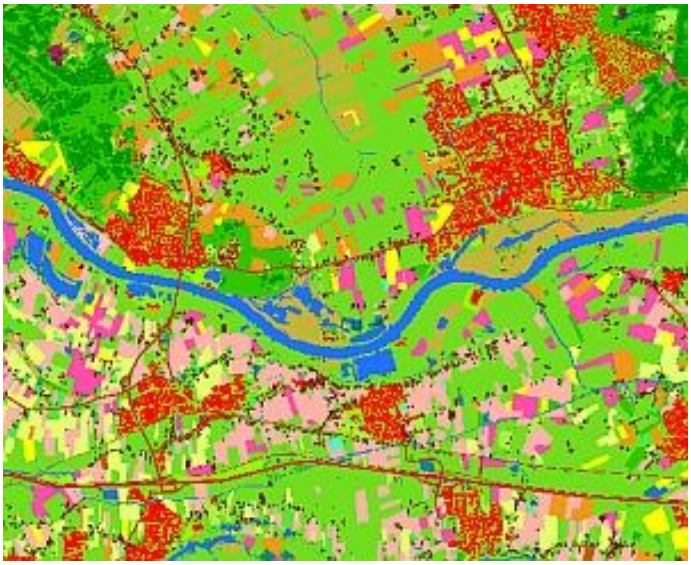


Figure 3.23: Extract from the LGN dataset (Alterra 2011).

3.5.2 Useful classes and geometry type

The LGN dataset is a grid dataset with a 25x25m resolution. The land use class given to an individual grid element refers to the dominating land use type present on this grid element. The LGN land use types which could be useful for the production of the culture code dataset, are presented in table 3.10 (in Dutch and translated into English).

Class	Class name (Dutch)	Class name (English)
1	Agrarisch gras	Agrarian grass
2	Mais	Corn/Maize
3	Aardappelen	Potatoes
4	Bieten	Beets
5	Granen	Cereals/Grain
6	Overige gewassen	Other crops
8	Glastuinbouw	Greenhouse growing
9	Boomgaarden	Orchards
61	Boomkwekerijen	Tree nurseries
62	Fruitkwekerijen	Fruit plantations

10	Bloembollen	Flower bulbs
11	Loofbos	Deciduous forest
12	Naaldbos	Conifer forest
16	Zoet water	Fresh water
17	Zout water	Salt water
20	Bos in primair bebouwd gebied	Forest in primarily built-up area
22	Bos in secundair bebouwd gebied	Forest in secundarily built-up area
23	Gras in primair bebouwd gebied	Grass in primarily built-up area
24	Kale grond in bebouwd gebied	Bare soil in built-up area
28	Gras in secundair bebouwd gebied	Grass in secundarily built-up area
30	Kwelders	Marshes
31	Open zand in kustgebied	Open sand in coastal area
32	Duinen met een lage vegetatie (<1m)	Dunes with a low vegetation (<1m)
33	Duinen met een hoge vegetatie (>1m)	Dunes with a high vegetation (>1m)
34	Duinheide	Dune moor
35	Open stuifzand en/of rivierzand	Open drift-sand and/or river sand
36	Heide	Moor
37	Matig vergraste heide	Moderately grassed moor
38	Sterk vergraste heide	Strongly grassed moor
39	Hoogveen	Peat moor
40	Bos in hoogveengebied	Forest in peat moor area
41	Overige moerasvegetatie	Other swamp vegetation
42	Rietvegetatie	Reed vegetation
43	Bos in moerasgebied	Forest in swamp area
45	Natuurgraslanden	Nature grass areas

Table 3.10: Object classes distinguished between in the LGN dataset (Alterra 2011).

3.5.3 Mappings with culture codes

Figure 3.24 shows the relationships between the LGN land use classes, presented in table 3.10, and the individual culture codes. The culture codes shown in red refer to culture codes which are no longer produced, but which could easily be produced by making use of the LGN dataset.

3.5.4 Quality and up-to-dateness

There is no strict up-date cycle for the LGN dataset. Every three till five years a new version of the LGN is produced. The most recent version of the dataset is the LGN6 version. This version dates from 2007/2008 (Alterra 2011).

According to Alterra (Alterra 2011), the quality of the LGN dataset is lower than that of the BRP dataset. Alterra has performed a quality analysis on their own dataset by comparing their attribute values (on type of crop, type of agricultural activity, etc.) to the ones stored in the BRP. They conclude that 84,8% of their data is right, meaning that this percentage is equal to the data found in the BRP dataset (Alterra 2011).

As the LGN dataset is a grid dataset, the positional accuracy/precision is not very high. The grids are 25x25m, so one could argue that the positional accuracy is also about 25m, which is not very good in comparison to the other candidate source datasets described and analyzed in this chapter.

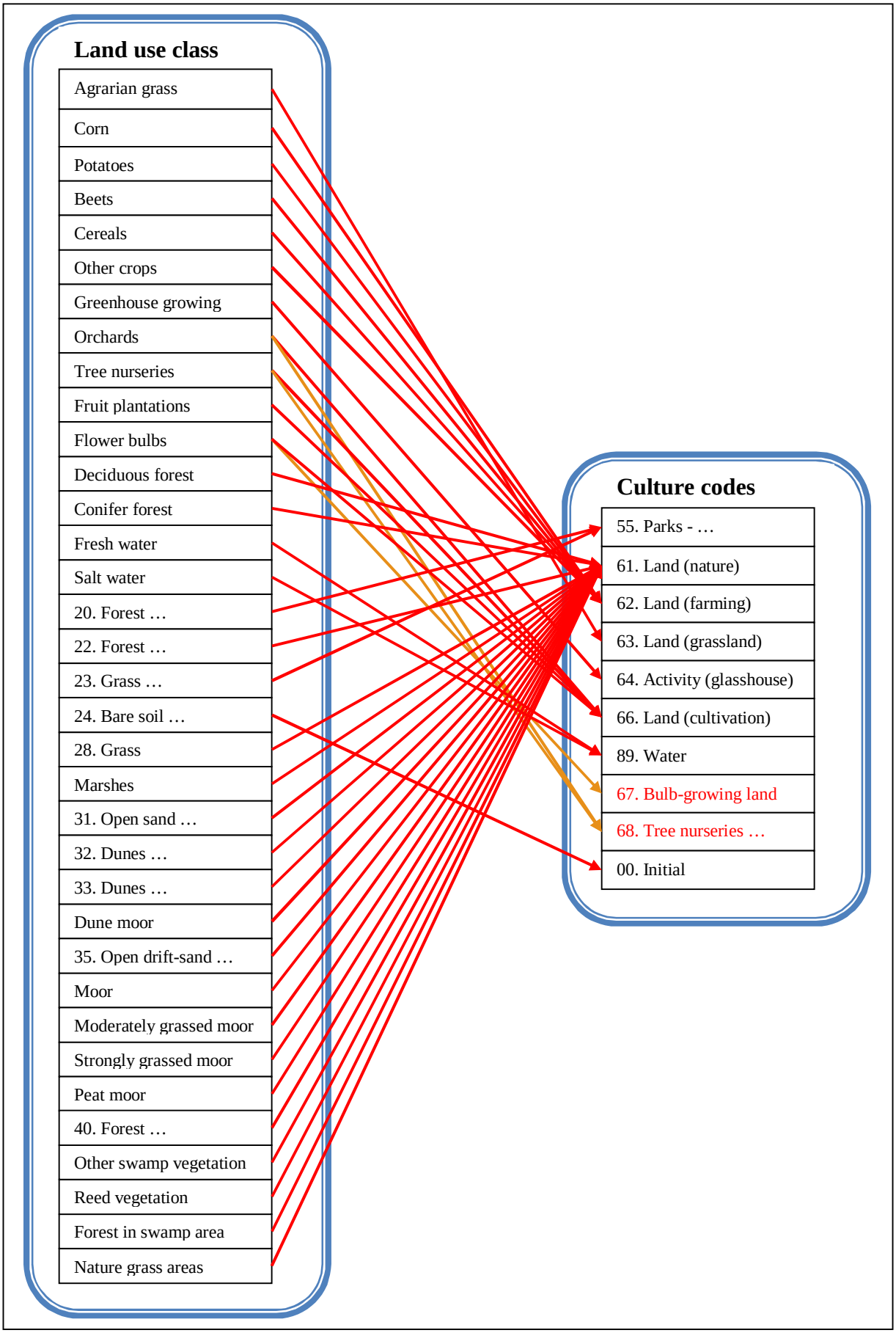


Figure 3.24: Mappings between LGN land use classes and culture codes.

3.5.5 Strengths and weaknesses

Table 3.11 presents the strengths and weaknesses of the LGN dataset.

Strengths	Weaknesses
+ Dataset covers the whole surface of the Netherlands	- High costs involved for obtaining the dataset.
+ Contains a large number of useful classes for the production of agriculture- or nature-related culture codes	- No strict update cycle for the dataset. Updated every three to five years.
	- According to Alterra, the LGN dataset has a lower quality than the BRP dataset.
	- Lower positional accuracy because the LGN dataset is a grid dataset.

Table 3.11: Strengths and weaknesses of the LGN dataset.

3.6 – AKR and LKI

The AKR and LKI datasets are produced and maintained by Kadaster. The LKI dataset contains geographical objects representing the cadastral parcels. One of its main functions is giving insight into the location, shape and size of cadastral parcels. Furthermore, the LKI contains the cadastral parcel numbers which can be used as entrance points for making use of administrative information from the cadastral administrative dataset (van den Brink 2010). This administrative information is part of the AKR dataset.

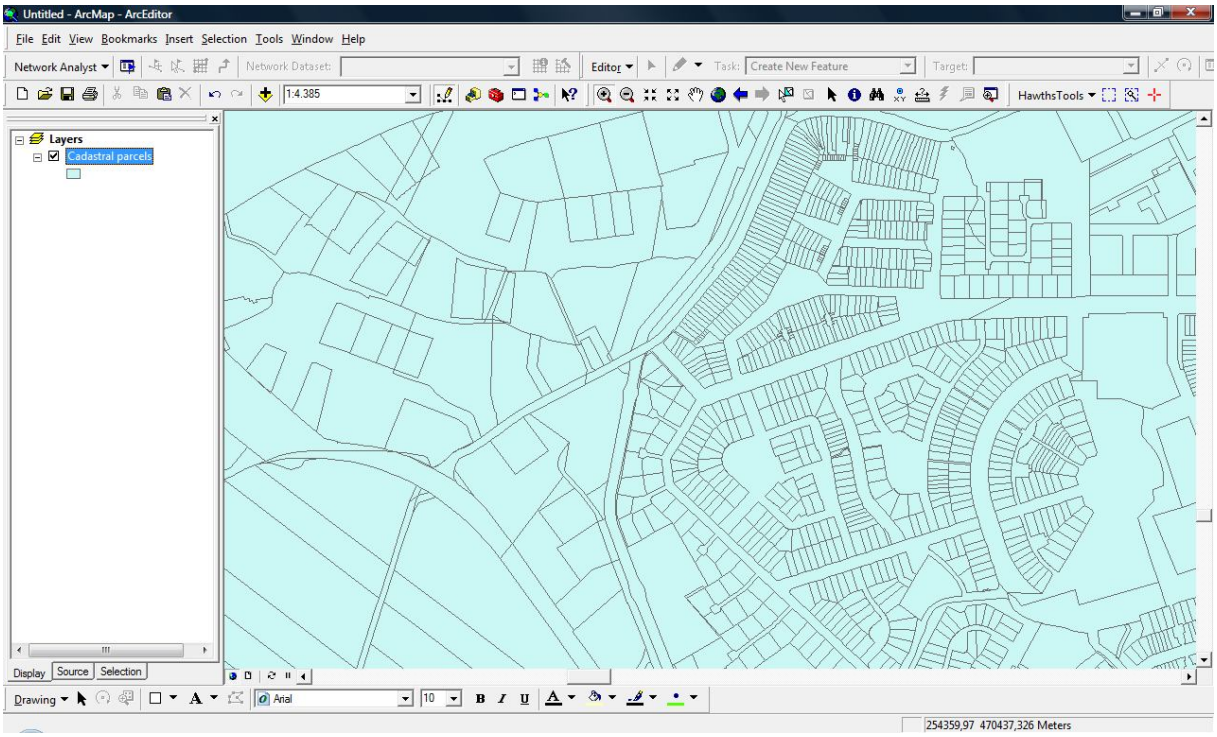


Figure 3.25: Visualization of cadastral parcels, part of the LKI cadastral parcel dataset in ArcGIS.

The culture codes are administrative attributes added to the cadastral parcel in the AKR dataset. Another type of information stored in the AKR dataset are the apartment rights (Kadaster 2011a). These apartment rights are useful information for the production of the

culture code “Residential (apartment)”. Combined with X and Y coordinates, the apartment rights dataset can be used in a GIS software package as point objects.

The LKI dataset is useful for the production of the culture code dataset, as it could serve to define the spatial entities for which the culture codes have to be calculated. Figure 3.25 visualizes the LKI dataset.

3.7 – Deciding on which source datasets to use

In this section it is decided which source datasets are used for the production of the culture codes in this research project. A number of criteria is taken into account when making this decision:

- The contents of the candidate source datasets. In order to be able to produce as many culture codes as possible with the selected source datasets, these datasets should contain most needed information.
- The quality of the candidate source datasets. If the quality of a dataset is very low or if other datasets – offering (almost) the same information – have a better quality, then the dataset is not selected for the production of the culture codes.
- The accessibility of the candidate source datasets. If a candidate source dataset is not available for the research, or when high costs are involved when obtaining the dataset, then this dataset is not selected for the production of the culture codes.

The AKR and LKI, BAG and TOP10NL datasets have been selected for the production of the culture codes in the next sections of this thesis. The LKI dataset is selected, as it contains the location, shape and size of the cadastral parcels for which the culture codes should be calculated. The AKR dataset contains information on the location of apartment rights, which is information not available in any other candidate source dataset. The reason for selecting the BAG dataset is the fact that it contains high quality building polygons and that is should have a high level of completeness being part of the Dutch system of key registers. Furthermore, the BAG dataset contains historical data, which makes it possible to calculate culture codes for a cadastral parcel at a certain time in history. It should, however, be noted that this is only possible for the time period since the creation of the BAG dataset (with the first municipalities completing the BAG dataset only a few years ago and the last one in spring 2011). Finally, the TOP10NL dataset is selected as it contains much detail on the attribute level. As a result, the TOP10NL object classes can be used for the production of a long list of culture codes (also visualized by the large amount of conceptual mappings). Like the BAG dataset, the TOP10NL dataset contains historical data. However, these historical data are not automatically delivered with the new version of the dataset, but is stored in a central database.

Another important reason for using these four datasets is their accessibility. The AKR, LKI and TOP10NL datasets are created and maintained by Kadaster, and can, therefore, easily be obtained and used for creating the culture codes. The BAG dataset is not produced by Kadaster, but the central database is maintained by and stored at Kadaster, making the data easily accessible. Furthermore, all datasets can be used by Kadaster without payments.

The other three datasets analyzed in this chapter (BRP, GBKN/BGT and LGN) are not used in the research. The BRP dataset is not used, as it is not currently not made available for third parties by the Ministry of Agriculture. The ministry in a transition phase when the BRP dataset is concerned. The BRP dataset could, however, in the future be useful for the production of agriculture- and nature-related culture codes, as the dataset is updated every year and contains detailed information on the type of agricultural usage.

The LGN is not used in the research project either, for a number of reasons. First, the LGN is a grid dataset, with every pixel containing the main land use type at that site. The positional accuracy is, therefore, not very high as local variation in land use cannot be presented in the LGN dataset. Second, the quality and up-to-dateness of the TOP10NL and BRP datasets are better than those of the LGN dataset. Third, for the production of the culture codes, the LGN dataset does not offer much extra information when comparing it with the TOP10NL dataset. The LGN dataset does offer some extra information (such as the location of greenhouses), but in general it offers more detailed information, which is not necessarily needed for the production of the culture codes. Fourth, obtaining the LGN dataset would cost at least €63.000, whereas the selected dataset can be obtained without payments.

Finally, the GBKN/BGT dataset is not used in the research. The most important reason for not using the GBKN dataset is the geometry type. The GBKN dataset consists of line elements, which makes it hard to use it for calculating percentages of the cadastral parcels covered by a certain land use type. The BGT will consist of polygons and, therefore, be a more useful dataset for the production of the culture code. However, as the BGT is not available yet for the whole surface of the Netherlands (only a small number of pilot/test municipalities has converted the GBKN into polygons) this is not a direct option for the production of the culture codes. Furthermore, the GBKN shows a large variation in quality (between urban and rural areas). The introduction of the BGT should, however, help to increase the quality of the large-scale topographical dataset. Therefore, it is suggested that Kadaster should follow the developments concerning the BGT dataset, as it might become a useful and high-quality dataset in the future.

Summing up, the AKR, LKI, BAG and TOP10NL datasets have been selected for the production of the culture codes in this research. This decision is based on the contents, quality and accessibility of the datasets. The author expects the datasets to offer useful information for the production of most current culture codes. Furthermore, the quality of the source datasets is relatively high and the datasets are easily accessible.

Chapter 4 – Exploring possibilities of using source datasets

In chapter 3, a description was given for the candidate source datasets for the Kadaster culture codes. Information was provided on the general content, production method, geometry, the attributes that might be suitable for the culture code derivation, the up-to-dateness and quality. This chapter goes a step further, by exploring the possibilities of using these candidate source datasets for the production of the culture code dataset. It does so by first overlaying the different datasets and presenting a number of possibilities and problems which arise when doing so (section 4.1). Then, in section 4.2, decision rules that should help to find the right culture code(s) for the cadastral parcels are being formulated. In section 4.3, the decision rules are integrated into a decision tree. The decision rules and tree will be tested in chapter 5.

4.1 – Overlay of candidate source datasets: possibilities and problems

In ArcGIS, an overlay of the candidate source datasets with the cadastral parcel dataset has been performed, in order to be able to visualize possibilities for the production of the culture codes. Sections 4.1.1 and 4.1.2 present examples of these overlays, respectively BAG buildings and residential objects (with purpose of use) in residential areas, and the TOP10NL terrain classes for non-built-up culture codes. Next to possibilities, the overlay procedure also reveals a number of problems to be overcome. Examples of such problems are: the presence of more than one purpose of use at a cadastral parcel (section 4.1.3), distinguishing between agricultural and industrial usage of a parcel (4.1.4), the presence of more than one non-built-up land use type at a cadastral parcel (4.1.5), and the presence of line- and/or point-elements on a parcel (4.1.6). Section 4.1.7 serves as a summary of section 4.1.

4.1.1 Possibility: BAG buildings and purpose of use in residential areas

Figure 4.1 shows the overlay of the BAG dataset with the cadastral parcels in a residential area. The orange area-shaped objects represent the buildings from the BAG dataset, whereas the green point-shaped objects visualize the residential units. The cadastral parcel boundaries are represented by the brown line-shaped objects in the figure.

The BAG buildings offer only geometric information, but do not contain information on the usage of these buildings. Information on the usage can, however, be obtained from the residential units dataset, a subset of the BAG dataset containing information on the purpose of use of these residential units. By combining the geometric information from the BAG buildings dataset with the attribute information from the BAG residential units dataset, it becomes possible to give information about the location and usage of a certain building at the same time.

In the case of figure 4.1, the map shows a residential area. As a result, all residential units located in this area have the purpose of use "residential". By performing a spatial join[9], the purposes of use can be added to the building polygons or the cadastral parcels. This would make it possible to assign a culture code "Residential" to those cadastral parcels which have a BAG building with the purpose of use "residential" located on them.

[9] A spatial join can be defined as "a type of table join operation in which fields from one layer's attribute table are appended to another layer's attribute table based on the relative locations of the features in the two layers" (ESRI 2009). In the example described in section 4.1.1 fields from the layer "purpose of use" are added to the layer "BAG buildings" when they are sharing a location on the map (in other words: when the point-shape purpose of use objects are located inside an area-shaped BAG building object).

In short, the BAG dataset offers useful information for assigning the culture code "Residential" to cadastral parcels. More general, the procedure described above can be applied to other purposes of use as well.

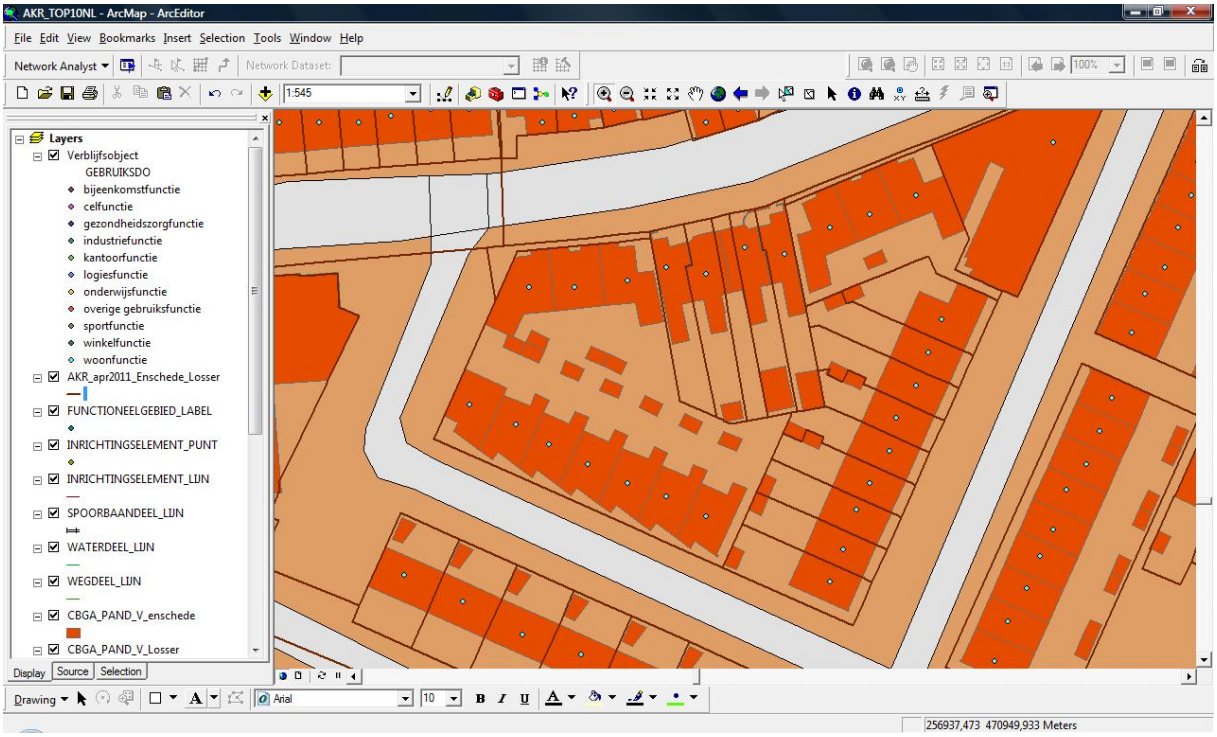


Figure 4.1: Result of overlaying BAG dataset (with buildings and purposes of use) and the cadastral parcel dataset. Source: datasets obtained from Kadaster.

4.1.2 Possibility: TOP10NL terrain classes for non-built-up culture codes

Section 4.1.1 showed the possibilities offered by the BAG dataset for defining the built-up culture code "Residential" in residential areas, with the method being applicable to other built-up culture codes too. For the non-built-up culture codes, e.g. water, roads, land (nature), land (farming) and land (grassland), the TOP10NL dataset could be used. As has been shown in section 3.3, the TOP10NL offers useful information for a wide range of culture codes. This section shows how the TOP10NL offers possibilities for determining the non-built-up objects located on cadastral parcels.

Figure 4.2 shows an extract from the TOP10NL and the cadastral parcel dataset. The figure shows that the TOP10NL dataset makes it possible to find one or more land use types for every cadastral parcel. In the extract the land use types water, road, forest, grassland, farming land, etc. can be found.

In short, the TOP10NL dataset could serve as a good counter-part of the BAG dataset. Whereas the BAG dataset could be used for built-up culture codes, the TOP10NL could be primarily[10] used for the non-built-up culture codes.

[10] Primarily, as the TOP10NL also offers information on types of buildings. These types of buildings are to be used as well and could help to overcome the limitations of the BAG in respect to the production of the culture codes.

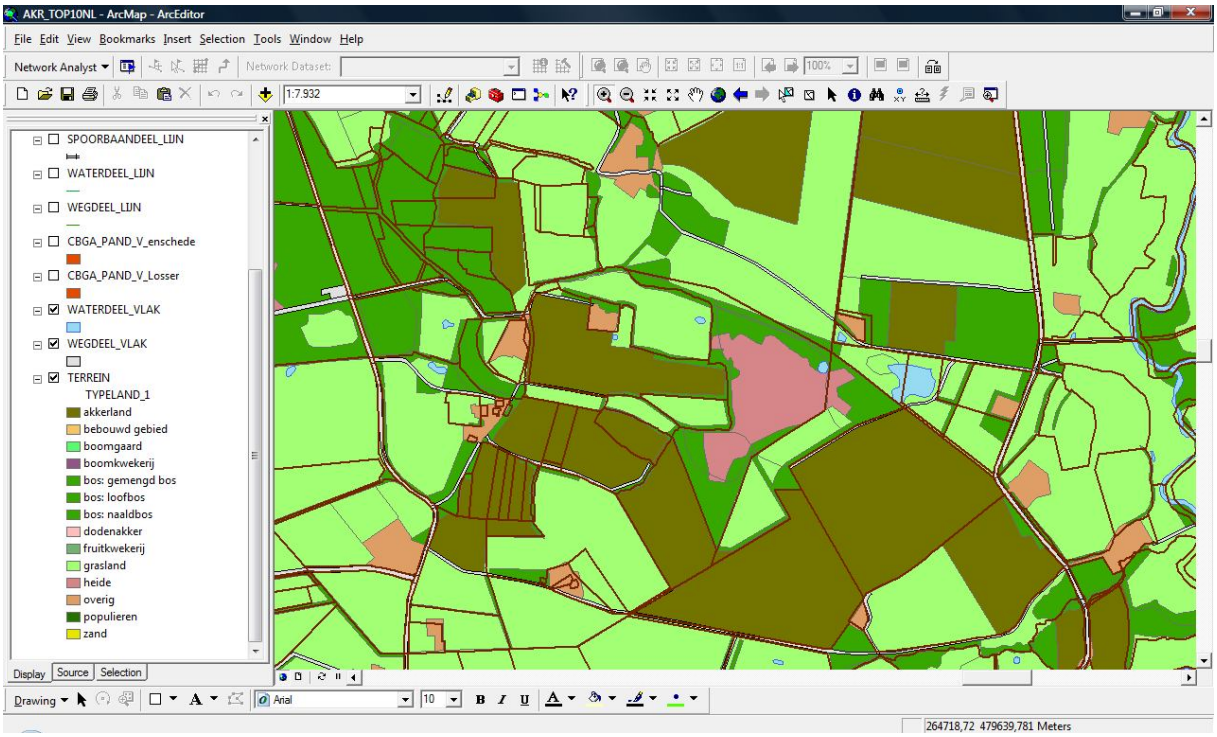


Figure 4.2: Result of overlaying TOP10NL (terrain, water objects and road sections) and the cadastral parcel dataset. Source: datasets obtained from Kadaster.

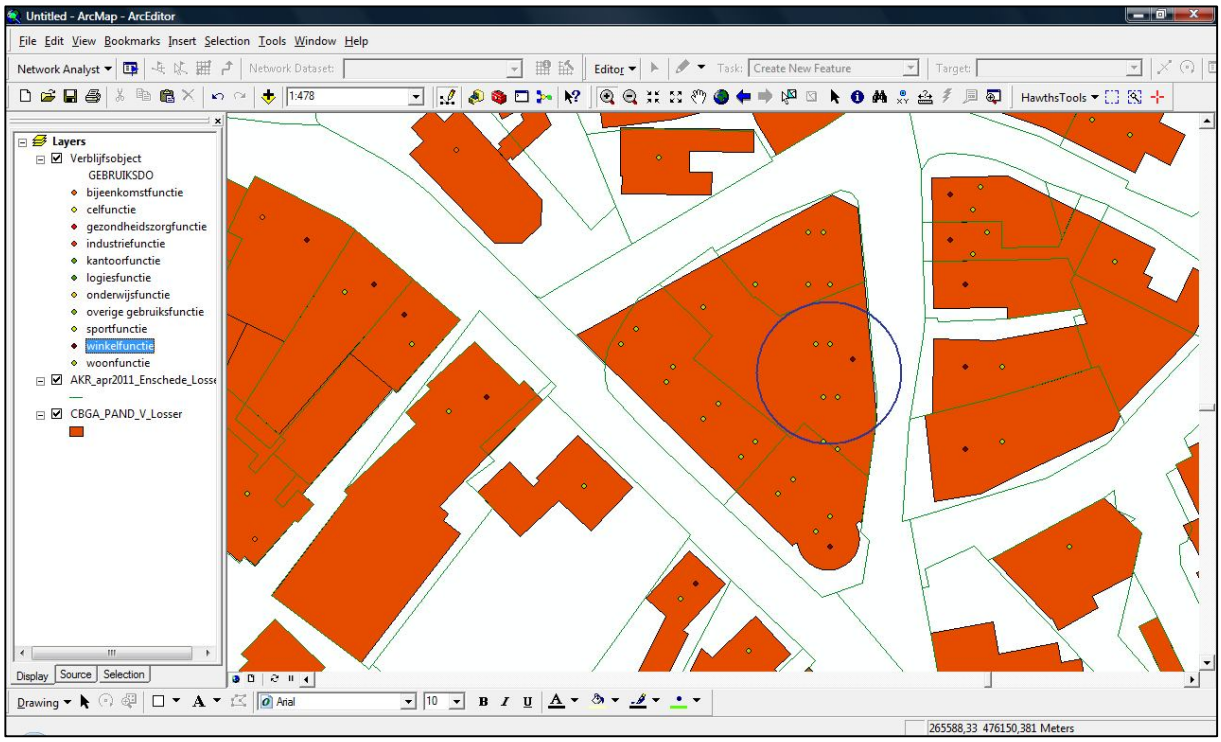


Figure 4.3: More than one purpose of use type located on one cadastral parcel (see blue circle). Source: datasets obtained from Kadaster.

4.1.3 Problem: more than one purpose of use per cadastral parcel

Section 4.1.1 showed that the BAG dataset, with its building polygons and purpose of use, offers interesting information for the production of the (built-up) culture codes. However, in figure 4.1 there was only one purpose of use type located on an individual cadastral parcel. In

many cases the situation is more complicated, with more than one purpose of use type located on a cadastral parcel. Figure 4.3 shows an example of this problem.

Such a situation is not always a problem, as in some cases one could assign the culture code "residential with activity" to a cadastral parcel with more than one purpose of use located on it. However, this is not possible when, for example, offices and industry are located on the same cadastral parcel. In such a case, one should give priority to one of these two purposes of use or, as has been applied in this thesis, the cadastral parcel could be given the culture code "Special properties".

4.1.4 Problem: agricultural vs. industry

A second problem arises when trying to make a difference between agriculture and industry. In the purposes of use for residential objects in the BAG dataset both land use types are referred to with the term "industry". In other words, the BAG dataset does not make a distinction between industry and agriculture. As the culture codes dataset does make a distinction between these two types of land use, a solution has to be found for this problem. Two solutions might be appropriate.

First, one could try to get additional information from the TOP10NL dataset. The TOP10NL dataset, as stated above, contains area-shaped land use objects (terrain). However, as illustrated by figure 4.4, the land use class, assigned to areas located underneath the BAG buildings with purpose of use "industrial", is "other". This is true for both agricultural and industrial locations. Therefore, it is not possible to use the TOP10NL to gather additional information for determining which BAG buildings are industrial and which are agricultural.

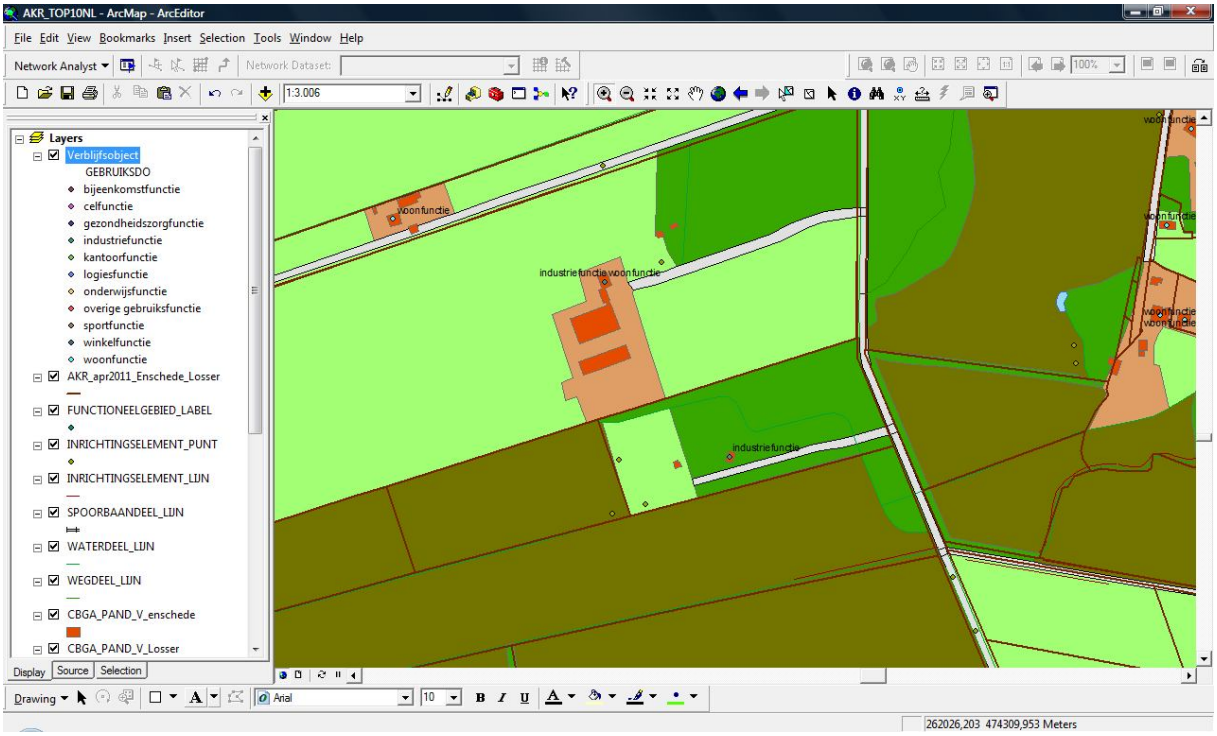


Figure 4.4: Industrial function in agricultural surroundings: industry or agriculture? Source: datasets obtained from Kadaster.

Second, one could work with neighborhood functions. For example, one could argue that BAG buildings with the purpose of use "industrial" which are within a distance of 50 meters from farming land, grassland or tree nurseries, should be assigned the culture code "activity (agricultural)". Vice versa, those objects not being within a distance of 50 meters from farming land, grassland or tree nurseries, should be assigned the culture code "activity

(industry)". Although there is a logical reasoning behind this second option, in some cases this could lead to the wrong culture code, as industrial companies can be located in areas with mainly agricultural activity, and agricultural companies can be located in an area without grassland, farming land or tree nurseries in the direct surroundings.

In short, the BAG dataset does not make it possible to distinguish between industry and agriculture. Two possible solutions have been elaborated in this section: adding information from the TOP10NL terrain objects and using neighborhood functions. The first solution cannot be applied, whereas the second could only be applied with severe limitations.

4.1.5: Problem: more than one land use type on one cadastral parcel

Another problem arises when more than one land use type is located on a single cadastral parcel. This is no problem as long as these land use types consist of one built-up and one non-built-up land use type – because then two culture codes can be assigned to the cadastral parcel. It becomes a problem when, for example, there are more than two land use types located on the individual parcel or when two built-up or two non-built-up elements are located on the same cadastral parcel. Two examples of this problem are the following.

First, in figure 4.5, the BAG element in the blue circle crosses the cadastral parcel boundary. A small part of the BAG building is located on a cadastral parcel covered to a large extent by road sections. One could assign two culture codes to this cadastral parcel: one for the BAG building (built-up code) and one for the road sections (non-built-up code). However, it is very well possible that the building is only located on it because of an error in the geometry of one of the source datasets, differences in positional accuracy, or differences in the production method. Therefore, it would be better to define a decision rule that makes sure such geometrical errors do not influence the assigning of culture codes to the cadastral parcels. One possible way of doing this, would be by working with percentages of the cadastral parcels covered by BAG building polygons. If, for example, less than 5% of the cadastral parcel is covered by a BAG building, then the parcel only gets a non-built-up culture code.

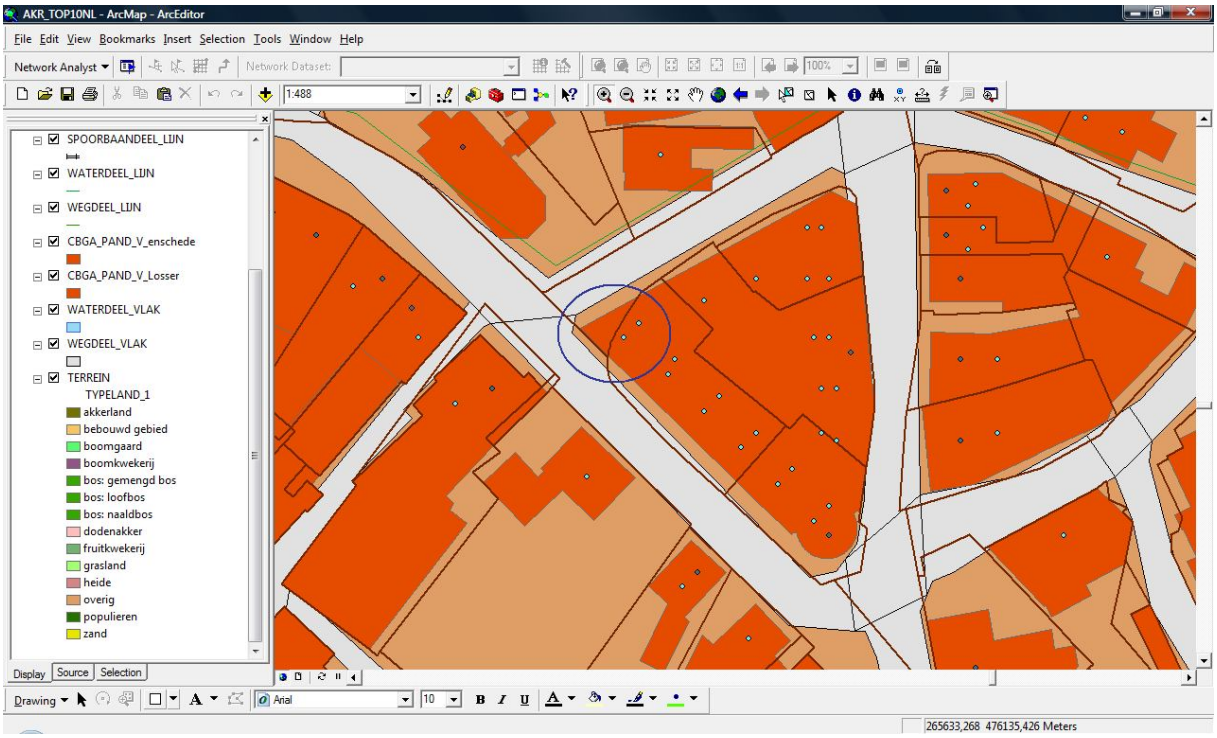


Figure 4.5: Part of a BAG building crossing the cadastral parcel boundary. Source: datasets obtained from Kadaster.

Again, this method could cause errors. Some residences in rural areas have large gardens, causing the percentage of the parcel covered by buildings to be very low. Still, one would like such a cadastral parcel to have both a built-up and a non-built-up culture code. Therefore, if chosen to apply percentages, it is interesting to study if different percentages fur rural and urban areas or different percentages for smaller and larger cadastral parcels will give better results.

Second, if more than one non-built-up element is located on a single cadastral parcel, it might be difficult to assign a single culture code to this parcel. An example is shown in figure 4.6. The picture shows a park, which consists of a number of cadastral parcels. The largest cadastral parcel contains a number of different non-built-up land use types: water, grassland, forest and roads. How to assign the correct culture code to this cadastral parcel?

One could apply a percentage method. Looking at figure 4.6, it is quite well possible that the land use type “grassland” would have the largest percentage covering the cadastral parcel. So, when using the percentage method, the cadastral parcel would get the culture code “land (grassland)”.

Another option would be to give priority to a certain land use type. For example, one could say that when roads are located on a certain parcel, and when it covers at least 5% of the parcel, the culture code “roads” should be assigned to such a parcel. In figure 4.6, it is well possible that the road sections cover more than 5% of the total parcel. As a result, when applying this priority rule, to the cadastral parcel the culture code “roads” would be assigned.

In both scenarios described in this section, one of the land use types present on the cadastral parcel has been assigned to the parcel as culture code. However, as stated in the introduction of this section, a park is located on the cadastral parcel, and one would, therefore, prefer to have the culture code “parks – public gardens” to be assigned to the cadastral parcel. This would make the procedure much more complex. A possibility would be to assign the culture code “parks – public gardens” to those cadastral parcels in urban areas which contain the land use types water, grassland and forest. By using this procedure, the cadastral parcel in figure 4.6 would be given the culture code “parks – public gardens”.

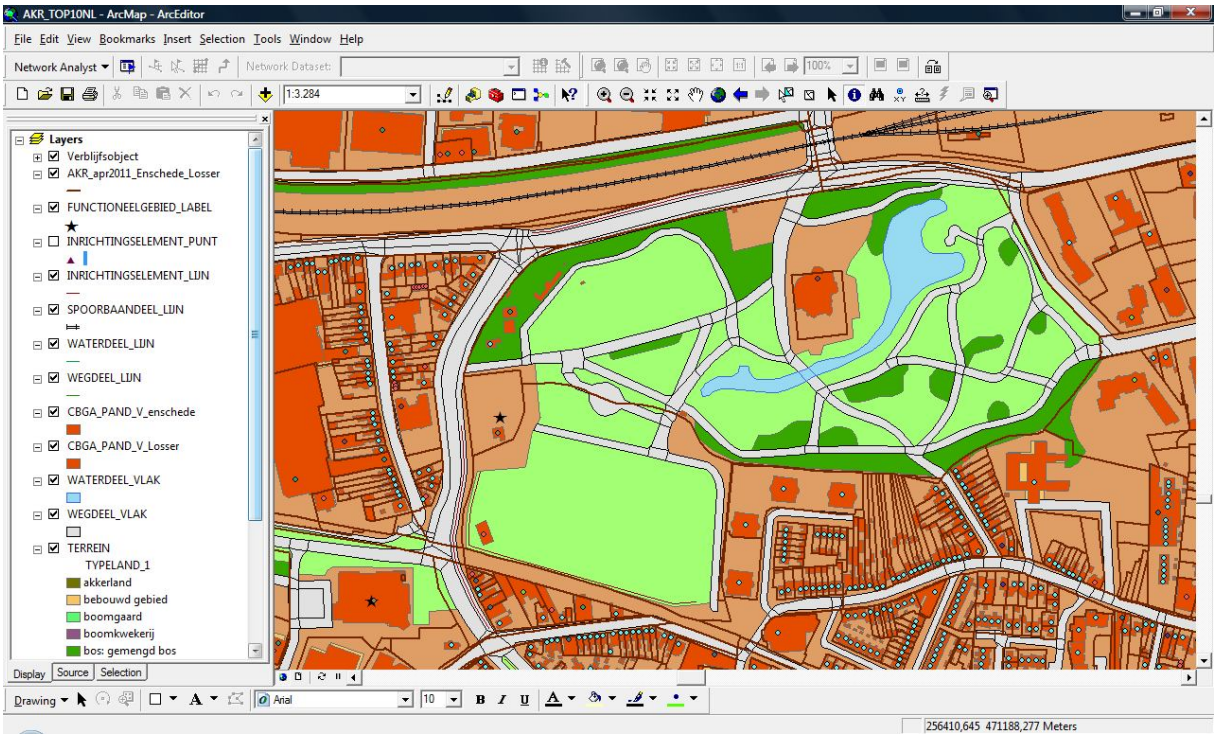


Figure 4.6: A park: how to determine the right culture code in case of many different land use types located on one cadastral parcel? Source: datasets obtained from Kadaster.

4.1.6 Problem: line- and point-shaped objects

A final problem to be discussed in this chapter, occurs when a cadastral parcels contains line- and/or point-shaped objects on the one hand, and area-shaped objects on the other hand. Figure 4.7 presents an example of such a problem. The cadastral parcel inside the blue box contains the area-shaped land use types forest, other and roads. However, the cadastral parcel does also contain a line-shaped element: a railroad. It is not possible to calculate percentages of the parcel covered by a line-shaped object (or a point-shaped object). Therefore, if working with the percentages method only, the cadastral parcel would never get the culture code "public transport", although this land use type might be the most important for the cadastral parcel. It is, therefore, necessary to look at another way of defining the culture codes when point- or line-shaped objects are located on a cadastral parcel.

A solution would be to work with priorities. For example, one could argue that cadastral parcels which contain a railroad section, should always get the culture code "public transport". In the case of the cadastral parcel in figure 4.7, this would mean the parcel would get the culture code "public transport", instead of "land (forest)" or "roads". It becomes even more complicated when different land use types with priority are located on the same cadastral parcel. In such a case, different priority levels should define which culture code is assigned. In other words: priority land use types with a higher priority level get priority over priority land use types with a lower priority level.

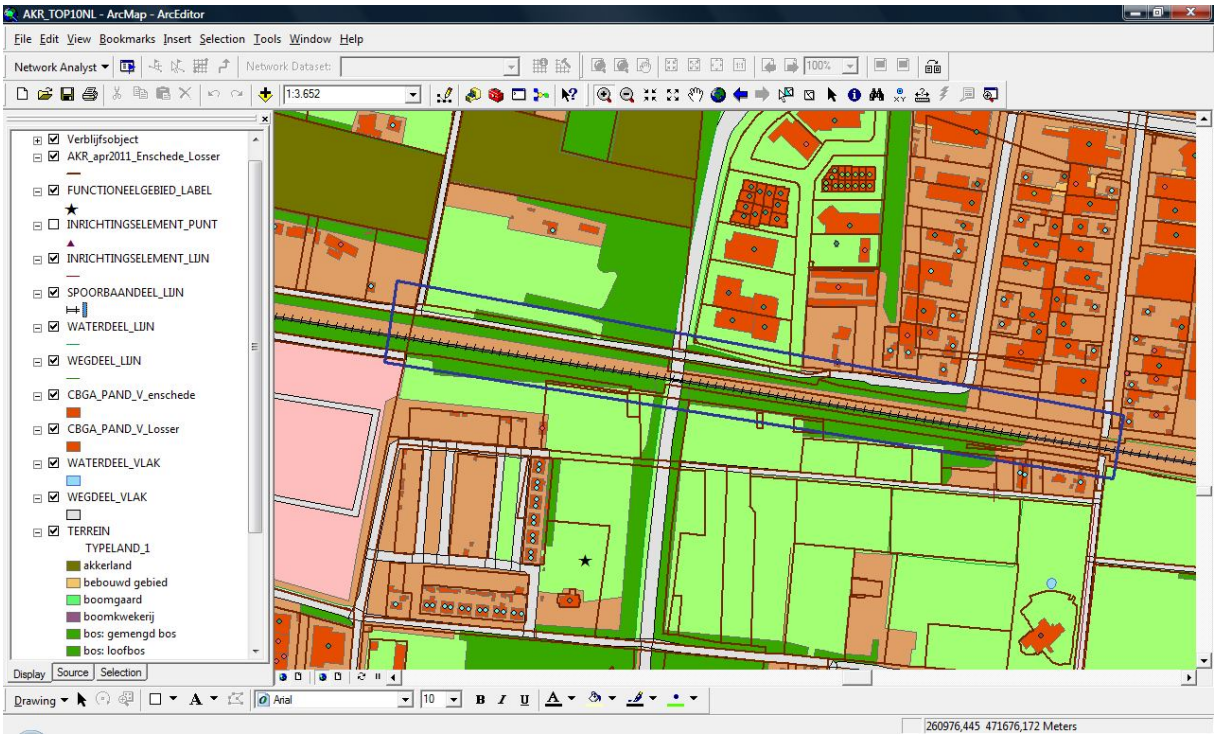


Figure 4.7: Example of a line-shaped object: a railroad. Source: datasets obtained from Kadaster.

4.1.7 Summary

The previous sections have presented a number of possibilities and problems concerning the production of a culture code dataset by making use of a number of source datasets. Some possibilities are the following:

1. The combination of BAG buildings and purposes of use offer the possibility to define culture codes ("residential") for cadastral parcels in residential areas;
2. The terrain land use types distinguished between in the TOP10NL offer the possibility to define non-built-up culture codes for many cadastral parcels.

Next to these possibilities, the following problems have been presented:

1. In some cases problems occur when more than one purpose of use (for the residential objects in the BAG dataset) are located on a single cadastral parcel;
2. It is hard to distinguish between agricultural and industrial built-up elements, as the BAG dataset does not make a distinction between them;
3. Problems occur when more than one land use type (e.g. more than one non-built-up land use type) is located on a single cadastral parcel;
4. It is not possible to work with a 'percentage' procedure when line- or point-shaped objects are located on a cadastral parcel.

The possibilities presented in this section, serve as input for the decision rules to be formulated in section 4.2. The problems explained here serve to give insight into the complexity involved when producing a culture code dataset by using a number of source datasets.

4.2 – Formulating decision rules

In this section, decision rules are formulated for the production of the culture codes. These decision rules should help to find the 'correct' culture code for cadastral parcels. The first sub-sections deal with individual or groups of decision rules, whereas the next sub-section (4.3) integrates all decision rules into a decision tree. This decision tree will be tested in the next chapter.

The decision rules are based on the content of the source datasets and built upon the conceptual mappings presented in chapter 3. At the moment, no documentation on the exact meaning of the culture codes is missing. Therefore, a definition list has been created by the author. This list – see appendix C – contains definitions that are partly definitions taken from the online dictionary *Cambridge Dictionaries Online* (Cambridge University Press 2011). The other part of the definition explains how this culture code is applied; it describes when a culture code is given to a cadastral parcel. This list with definitions is, together with the content input presented in chapter 3, used for the formulation of the decision rules. All possible valuable attributes and object classes offered by the source datasets are used in the decision rules and no priorities are given to particular decision rules over others. Priorities are first applied in the decision trees, and are, therefore, explained in section 4.3.

4.2.1 Built-up vs. non-built-up

As referred to in section 4.1, the BAG dataset could be used to determine whether a cadastral parcel should get a built-up culture code or not. For making this decision, percentages of the parcels covered by BAG building polygons could be used. This procedure does not lead directly to a culture code, but rather limits the number of possible culture codes in those cases where only a non-built-up or only a built-up culture code can be assigned to the parcel. This step can, however, be used to define the "*bebouwingscode*" (see section 1.2) of a cadastral parcel. More precise, one could assign the following *bebouwingscodes* to the cadastral parcels[11]:

[11] In section 1.2 the list of possible *bebouwingscodes* also included the option "Special properties". In this thesis it is chosen to deal with "Special properties" as one culture codes, as it is also included in the list of culture codes (see appendix A).

- Built-up;
- Non-built-up;
- Non-built-up and built-up.

One could say that parcels completely covered by BAG building polygons should be assigned the code “built-up”, those parcels not containing BAG buildings the code “non-built-up”, and those partly covered by BAG polygons the code “non-built-up and built-up”. However, as has been argued in section 4.1.5, this causes inaccuracies in the source datasets to influence the culture code outcomes. As a result, it is proposed to use other percentages, which should help to decrease the impact of inaccuracies on the outcome of the production method. The following decision rules, and the percentages mentioned in these decision rules, are proposed in this document:

Decision rule 1: If more than 90% of a parcel is covered by BAG building polygons, then this parcel should get only a built-up culture code.

Decision rule 2: If less than 2% of a parcel is covered by BAG building polygons, then this parcel should get only a non-built-up culture code.

Decision rule 3: If 2% or more and 90% or less of a parcel is covered by BAG building polygons, then this parcel should get both a built-up and a non-built-up culture code.

The percentages used in the decision rules are not based on calculations, but are also not chosen completely arbitrarily. As stated in section 4.1.5, in urban areas it would be possible to work with a higher percentage than the 2% used in decision rule 2. However, in rural areas some residences have large gardens. Working with a percentage of, for example, 5% would cause these parcels to get only a non-built-up culture code, although the residences located on them are at least just as important as the gardens. Therefore, it has been chosen to use the 2% break. The proof of concept in the next chapter will prove whether this percentage is chosen well or if another percentage might give better results.

The 90% break between built-up and a combination of built-up and non-built-up also has to be tested during the proof of concept. The reasoning behind this percentage, is that the buildings covering more than 90% are thought to be more important than the less than 10% covered by non-built-up elements. Therefore, if more than 90% of a cadastral parcel is covered by BAG building polygons, this parcel should only get a built-up culture code.

4.2.2 Built-up culture codes

In the previous section, decision rules have been presented which make it possible to distinguish between cadastral parcels with only built-up elements, with only non-built-up elements and with a combination of built-up and non-built-up elements. The current section deals with these cadastral parcels having built-up elements and gives information on how to decide on which built-up culture code should be assigned to these parcels. It does so by describing the process of culture code production for the built-up culture codes, in the same order as they have been added to the list in appendix A (culture codes currently present in the cadastral database).

11 Residential

In order to create the culture codes for cadastral parcels having residential units located on them, the BAG dataset is used. More specific, the purpose of use “residential function” from

the residential objects sub-dataset is used. When a cadastral parcel contains built-up elements, and when "residential function" is the only purpose of use located on the parcel, then the culture code should be "residential".[12]

Decision rule 4: If only the purpose of use "residential function" from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code "Residential".

12 Residential (apartment)

An exception has to be made to decision rule 4 when apartments are located on the cadastral parcel, in which case the parcel should get the culture code "Residential (apartment)". These apartment rights are registered in the AKR dataset. It could also be argued that one gives the culture code "Residential (apartment)" to all cadastral parcels containing more than one residential object but only the purpose of use "residential function". However, it is possible that a parcel contains more than one building, with all separate buildings having an individual purpose of use object and not being an apartment (but, for example, separate houses). Therefore, the apartment rights give more reliable information for the production of the culture code "Residential (apartment)".

Decision rule 5: If apartment rights from the BRK dataset are located on a cadastral parcel containing only the purpose of use "residential function" from the BAG dataset, then this parcel should get the culture code "Residential (apartment)".

14 Residential (agricultural)

The production of the right culture codes becomes more complicated in case of the culture code "Residential (agricultural)". As has been explained in section 4.1.4, the BAG dataset does not make a distinction between industrial and agricultural. Both industrial and agricultural use of a building are referred to in the BAG dataset with the purpose of use "industrial function". The TOP10NL and BRP datasets do neither directly provide useful information; in the TOP10NL dataset the land use class located underneath the BAG buildings with the purpose of use "industrial function" is always "other", whereas the BRP does not contain data on the areas located underneath these buildings.

In section 4.1.4 it has been suggested already that it might be possible to work with distances to agricultural land, by making use of the land use types distinguished between in the TOP10NL dataset. Agricultural land includes, in this case, the TOP10NL land use classes farming land, orchard, tree nursery, fruit nursery and grassland. If the residential objects from the BAG dataset are within a distance of 50 meters from such a land use class, then the cadastral parcels should get the culture code "Residential (agricultural)". This is formulated in the following decision rule:

[12] In the decision rule it is not mentioned that it only counts for cadastral parcels which should get a built-up culture code. As section 4.2.2 only deals with the built-up culture codes, it is pre-assumed that all decision rules count only for cadastral parcels which should receive a built-up culture code.

Decision rule 6: If a cadastral parcel contains both the purpose of use "residential function" and the purpose of use "industrial function" from the BAG dataset, then this parcel should get the culture code "Residential (agricultural)", but only when the residential object is located within a distance of 50 metres from one of the following TOP10NL land use classes:

Farming land;
Orchard;
Tree nursery;
Fruit nursery;
Grassland.

18 Storeroom – shelter (garage-shed)

The datasets used in this research project do not offer much information on the presence of storerooms or shelters on cadastral parcels. In the BAG dataset they are part of the purpose of use "other function". The only information which could help to create this culture code is provided by the TOP10NL dataset: the type of building "coach-house". This attribute could be used, but the level of completeness of this culture code will be low.

Decision rule 7: If the type of building "coach-house" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Storeroom – shelter (garage-shed)".

21 Activity (office)

The information needed to create this culture code can be obtained from the BAG dataset, as this dataset contains the purpose of use "office function". If this is the only purpose of use located on a cadastral parcel, then this cadastral parcel should get the culture code "Activity (office)".

Decision rule 8: If only the purpose of use "office function" from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code "Activity (office)".

22 Activity (industry)

For the production of this culture code decision rule 6 can be re-used. However, in this case the culture code "Activity (industry)" should be assigned to the cadastral parcel when the parcel contains only the purpose of use "industrial function" and when the residential object is not within a distance of 50 meters from agricultural land.

Decision rule 9: If a cadastral parcel contains the purpose of use "industrial function" from the BAG dataset, then this parcel should get the culture code "Activity (industry)", but only when the residential object is NOT located within a distance of 50 metres from one of the following TOP10NL land use classes:

Farming land;
Orchard;
Tree nursery;
Fruit nursery;
Grassland.

23 Activity (utility)
Only the TOP10NL dataset offers information on utility artefacts. The following object classes offer attributes which could help to produce the culture code:

1. Functional area: the attribute “water treatment”;
2. Cultural landscape object: the attribute “gas extraction”;
3. Type of building: the attribute “powerplant” and “nuclear powerplant”.

Decision rule 10: If the functional area “water treatment” from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Activity (utility)”.

Decision rule 11: If the cultural landscape object “gas extraction” from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Activity (utility)”.

Decision rule 12: If one of the building types “powerplant” or “nuclear powerplant” from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Activity (utility)”.

27 Activity (retail trade)
For the production of this culture code, again the BAG dataset can be used, as this dataset includes the purpose of use “retail function”. If only this purpose of use is located on a cadastral parcel, then this parcel should get the culture code “Activity (retail trade)”.

Decision rule 13: If only the purpose of use “retail function” from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code “Activity (retail trade)”.

28 Activity (catering)
With the datasets used in this research project it is not possible to produce this culture code. The BAG dataset does not give enough information, as buildings with catering function are within the purpose of use class “gathering function”, which also contains other functions such as religious or cultural functions. Therefore, for the production of this culture code, additional information would be needed. The only attribute from the candidate source datasets offering useful information, is the type of building “road restaurant” from the TOP10NL dataset. If this building type is located on a cadastral parcel, this parcel should get the culture code “Activity (catering)”.

Decision rule 14: If the type of building “road restaurant” from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Activity (catering)”.

29 Activity (agricultural)
For the production of this culture code decision rule 6 can be re-used. However, in this case the culture code “Activity (industry)” should be assigned to the cadastral parcel when the parcel contains only the purpose of use “industrial function” and when the residential object is within a distance of 50 meters from agricultural land.

Decision rule 15: If a cadastral parcel contains the purpose of use "industrial function" from the BAG dataset, then this parcel should get the culture code "Activity (agricultural)", but only when the residential object is located within a distance of 50 metres from one of the following TOP10NL land use classes:

Farming land;
Orchard;
Tree nursery;
Fruit nursery;
Grassland.

34 Parking

Parking can be both a built-up (parking garages) as a non-built-up (parking lots) element. To be able to produce the built-up culture code "Parking", the TOP10NL object class "type of building" can be used, as it contains the attribute "parking garage".

Decision rule 16: If the type of building "parking garage" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Parking".

35 Defense

Like parking, the culture code "Defense" can be both built-up and non-built-up. To produce the built-up culture code "Defense" the attribute "military building" from the TOP10NL object class "type of building" can be used.

Decision rule 17: If the type of building "military building" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Defense".

37 Residential with activity

The culture code "Residential with activity" should be assigned to cadastral parcels which contain at least two types of purpose of use, with one being "residential function". The other purpose of use could be one of the following: retail trade, gathering, accommodation, industry or office function. As "Residential (agricultural)" has been defined above, in case of the purpose of use "industry", only those residential objects that are not within a distance of 50 meters from agricultural land should be used to produce the culture code "Residential with activity". As a result, this purpose of use gets a special decision rule, based on decision rule 6.

Decision rule 18: If a cadastral parcel contains the purpose of use "residential function" from the BAG dataset, then this parcel should get the culture code "Residential with activity" when the cadastral parcel does also contain one of the following purposes of use:

Retail function;
Gathering function;
Accommodation function;
Office function.

Decision rule 19: If a cadastral parcel contains both the purpose of use "residential function" and the purpose of use "industrial function" from the BAG dataset, then this parcel should get the culture code "Residential with activtity" when the residential object is NOT located within a distance of 50 metres from one of the following TOP10NL land use classes:

Farming land;
Orchard;
Tree nursery;
Fruit nursery;
Grassland.

44 Public transport

For the production of the built-up culture code "Public transport", which can also be a non-built-up culture code, the TOP10NL dataset can be used. The cultural landscape objects class contains the attribute "station" and the type of building object class contains the attributes "metro station" and "train station".

Decision rule 20: If the cultural landscape object "station" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Public transport".

Decision rule 21: If one of the building types "metro station" or "train station" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Public transport".

46 Air traffic

Also for the production of the culture code "Air traffic" the TOP10NL dataset can be used. In this case, the cultural landscape objects class contains the attribute "helicopter platform" and the type of building object class contains the attribute "traffic tower".

Decision rule 22: If the cultural landscape object "helicopter platform" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Air traffic".

Decision rule 23: If the building type "traffic tower" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Air traffic".

51 Recreation – sport

For the production of this built-up culture code, again the BAG dataset can be used, as this dataset contains the purpose of use "sports function". If a residential object with this purpose of use is located on a cadastral parcel, the parcel should get the culture code "Recreation – sport". Next to the BAG dataset, the object class type of building of the TOP10NL dataset contains useful information for the production of this culture code. The following attributes could be used to create the culture code "Recreation – sport": artificial ice track, horse riding school, recreation centre, sports building, stadium and swimming pool.

Decision rule 24: If only the purpose of use "recreational function" from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code "Recreation – sport".

Decision rule 25: If one of the following building types from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Recreation – sport":

- Artificial ice track;
- Horse riding school;
- Recreation centre;
- Sports building;
- Stadium;
- Swimming pool.

53 Residential (recreation)

The BAG dataset also contains a purpose of use "accommodation function", which could be used for producing the culture code "Residential (recreational)". If a residential object with the purpose of use "accommodation function" is located on a cadastral parcel, this parcel should be assigned this culture code.

Decision rule 26: If only the purpose of use "accommodation function" from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code "Residential (recreation)".

64 Activity (glasshouse)

Glasshouse activity cannot be identified by making use of the BAG and TOP10NL dataset. However, the LGN dataset (described in section 3.5, but not available for usage in this research) does contain this information. Furthermore, it might be possible that the BRP dataset will contain the same information. As both datasets are not available for this research project, a decision rule formulated in this section could not be tested. As a result, no decision rule is formulated.

72 Health

The culture code "Health" can be produced by making use of the purposes of use of the BAG dataset, as it contains a purpose of use "health function". If this purpose of use is located on a cadastral parcel, then this parcel should get the culture code "Health".

Decision rule 27: If only the purpose of use "health function" from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code "Health".

74 Education

Just like the culture code "Health", the culture code "Education" can be produced by making use of the purpose of use in the BAG dataset. The dataset contains the purpose of use "educational function". If this purpose of use is located on a cadastral parcel, the parcel should get the culture code "Education".

Decision rule 28: If only the purpose of use "educational function" from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code "Education".

76 Culture

The BAG dataset cannot be used to produce this culture code, as the cultural buildings are placed within the purpose of use class "gathering function", which does also include other usage types (e.g. religion). As a result, other sources need to be used for the production of this culture code. The type of building object class of the TOP10NL dataset does include useful information, as it contains the attributes "visitor centre" and "museum". These two attributes could be used to create the culture code "Culture".

Decision rule 29: If one of the building types "visitor centre" or "museum" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Culture".

77 Religion

Just like the cultural buildings, religious buildings are, in the BAG dataset, placed in the purpose of use class "gathering function". This means the BAG dataset cannot be used to produce the culture code "Religion". In this case too, the TOP10NL dataset offers useful information. The cultural landscape object class contains the attribute chapel. Furthermore, the object class type of building contains the attributes "chapel", "church", "monastery" and "religious building". Using these attributes, the culture code "Religion" can be produced.

Decision rule 30: If the cultural landscape object "chapel" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Religion".

Decision rule 31: If one of the building types "chapel", "church", "monastery" or "religious building" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Religion".

78 Police – Fire service

In the BAG dataset, police and fire service offices/buildings are not defined. They are part of the attribute class "other function". As a result, the BAG dataset cannot be used to assign this culture code to cadastral parcels. The type of building "police office" from the TOP10NL dataset can be used to produce the culture code. No information is available in the TOP10NL dataset about the location of fire service buildings.

Decision rule 32: If the building type "police office" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Police – Fire service".

79 Justice

Both the BAG dataset and the TOP10NL dataset offer information for the production of the culture code "Justice". The BAG dataset contains the purpose of use "prison/cell function", whereas the TOP10NL dataset contains the type of building attribute "prison".

Decision rule 33: If only the purpose of use “prison/cell function” from the BAG dataset is located on a cadastral parcel, then this parcel should get the culture code “Justice”.

Decision rule 34: If the building type “prison” from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Justice”.

85 Port

Only the TOP10NL dataset offers information which can be used for the production of the culture code “Port”. As part of the TOP10NL, the type of building object class contains the attributes “dock”, “lifeboat building” and “shipyard”.

Decision rule 35: If one of the building types “dock”, “lifeboat building” or “shipyard” from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Port”.

92 Land (new construction activity)

For the production of this culture code the status of buildings from the BAG dataset is needed. When the status of such a building is “building process started” or “building permit granted”, then this site is a new construction site. This means that if a culture code has been assigned to a cadastral parcel containing a building with one of these two statuses, then the cadastral parcel should actually get the culture code “Land (new construction activity)”. This counts for the following built-up culture codes:

- Activity (office);
- Activity (industry);
- Activity (utility);
- Activity (retail trade);
- Activity (catering);
- Activity (agricultural);
- Activity (glasshouse).

Decision rule 36: If a cadastral parcel contains a BAG building with the status “building process started” or “building permit granted”, then this cadastral parcel should get the culture code “Land (new construction activity)” when one of the following built-up culture codes has been assigned to the parcel:

- Activity (office);
- Activity (industry);
- Activity (utility);
- Activity (retail trade);
- Activity (catering);
- Activity (agricultural);
- Activity (glasshouse).

93 Land (new construction residential)

The same as has been mentioned for decision rule 35, counts for cadastral parcels containing buildings with a residential function and having the status “building process started” or

"building permit granted". These parcels should get the culture code "Land (new construction residential)" instead of the culture codes:

- Residential;
- Residential (apartment);
- Residential (agricultural);
- Residential with activity;
- Residential (recreation).

Decision rule 37: If a cadastral parcel contains a BAG building with the status "building process started" or "building permit granted", then this cadastral parcel should get the culture code "Land (new construction residential)" when one of the following built-up culture codes has been assigned to the parcel:

- Residential;
- Residential (apartment);
- Residential (agricultural);
- Residential with activity;
- Residential (recreation).

94 Funeral service

The only useful attribute class can be found in the type of building object class from the TOP10NL dataset: the "crematory" attribute class. When this type of building is located on a cadastral parcel, the parcel should get the culture code "Funeral service".

Decision rule 38: If the building type "crematory" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Funeral service".

4.2.3 Non-built-up culture codes

In section 4.2.2 the decision rules for the production of the built-up culture codes have been formulated. In the current section those decision rules which should assist in assigning culture codes to those cadastral parcels requiring a non-built-up culture code (see section 4.2.1). Again, this is done in the order of appearance in the list of culture codes presented in appendix A.

25 Land (industry)

None of the candidate source datasets offers much information on the presence of industrial land on cadastral parcels. Only the TOP10NL offers some direct information. This information appears as the functional area attribute "mine"; if a mine is located on a certain cadastral parcel, then this parcel should get the culture code "Land (industry)". Indirectly, the BAG dataset offers information for the production of this culture code. It is assumed in this research project, that to cadastral parcels containing industrial buildings (this means, parcels with the built-up culture code "Activity (industry)") which should also get a non-built-up culture code, the culture code "Land (industry)" is assigned.

Decision rule 39: If the functional area type "mine" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Land (industry)".

Decision rule 40: If the built-up culture code "Activity (industry)" is located on a cadastral parcel, and when this parcel should also get a non-built-up culture code, then this cadastral parcel should get the culture code "Land (industry)".

34 Parking

The road section object lass from the TOP10NL dataset offers useful information for the production of this culture code, as it contains the attributes "parking", "park + ride" and "carpool site". Cadastral parcels containing at least one of these attributes should be assigned the culture code "Parking".

Decision rule 41: If one of the main use types "parking", "park + ride" or "carpool site" from the TOP10NL road section object class is located on a cadastral parcel, then this parcel should get the culture code "Parking".

35 Defense

The TOP10NL contains the functional area type "military area". If this functional area type is located on a cadastral parcel, then this parcel should get the culture code "Defense". Furthermore, if a cadastral parcel has been given the built-up culture code "Defense" (see section 4.2.2), and if there is surrounding terrain located on the same parcel, then this parcel should also get the non-built-up culture code "Defense".

Decision rule 42: If the functional area type "military area" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Defense".

Decision rule 43: If the built-up culture code "Defense" is located on a cadastral parcel, and when this parcel also contains surrounding terrain, then this cadastral parcel should get the non-built-up culture code "Defense".

41 Roads

The TOP10NL contains the object class "road section". This object class contains a wide range of attributes, with the attribute "main use" being the most useful one for the production of the culture code "Roads". If the main use is either "fast traffic" or "mixed traffic", then the cadastral parcel should get the culture code "Roads". This applies only to pave roads, so non-paved roads have to be excluded.

Decision rule 44: If a cadastral parcel contains a road section with either main use type "fast traffic" or main use type "mixed traffic" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Roads".

44 Public transport

Again, the road section object class from the TOP10NL dataset offers useful information, as it contains a main use type "bus traffic". Furthermore, the TOP10NL dataset contains the object class "railroad section". If a railroad section (which can be a metro, tram, train or mixed

railroad) is located on a cadastral parcel, then this parcel should also get the culture code "Public transport".

Decision rule 45: If a cadastral parcel contains a road section with main use type "bus traffic" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Public transport".

Decision rule 46: If a cadastral parcel contains a railroad section from the TOP10NL dataset, then this cadastral parcel should get the culture code "Public transport".

46 Air traffic

Two object classes from the TOP10NL dataset offer information which can be used for the production of the culture code "Air traffic". On the one hand, the functional area object class contains the attributes "helicopter landing site" and "airport". On the other hand, the road section object class contains the main use type "air traffic" (non-paved roads should be excluded). If one of these attributes from the TOP10NL dataset is found on a cadastral parcel, then this parcel should get the culture code "Air traffic".

Decision rule 47: If one of the functional area types "helicopter landing site" or "airport" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Air traffic".

Decision rule 48: If a cadastral parcel contains a road section with main use type "air traffic" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Air traffic".

47 Pipes – Tubes

The candidate source datasets used and described in this research project do not contain information which could be used for the production of this culture code. As a result, additional datasets/information is required for the production of the culture code "Pipes – Tubes".

51 Recreation – Sport

Four object classes from the TOP10NL dataset offer information for the production of the culture code "Recreation – Sport". These four object classes and the interesting attributes are the following:

1. Road section – main use: horse riding;
2. Water object – function: natural pool and swimming pool;
3. Cultural landscape object: shooting range;
4. Functional area: bungalow park, camping place, caravan park, circuit, crossing track, zoo, golf course, karts track, recreational area, hippodrome, ski slope, sports site, tennis course and ice track.

Decision rule 49: If a cadastral parcel contains a road section with main use type "horse riding" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Recreation – Sport".

Decision rule 50: If a cadastral parcel contains a water object with one of the functions "natural pool" or "swimming pool" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Recreation – Sport".

Decision rule 51: If the cultural landscape object "shooting range" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Recreation – Sport".

Decision rule 52: If one of the following functional area types from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Recreation – sport":

- Bungalow park;
- Camping place;
- Caravan park;
- Circuit;
- Crossing track;
- Zoo;
- Golf course;
- Karts track;
- Recreational area;
- Hippodrome;
- Ski slope;
- Sports site;
- Tennis course;
- Ice track.

55 Parks – Public gardens

As has been argued in section 4.1.5, it is hard to define which cadastral parcels should get the culture code "Parks – Public gardens" by making use of the terrain classes distinguished between in the TOP10NL dataset. One could assign the culture code to cadastral parcels having a combination of grassland, forest and water objects. However, this is no 'safe' method, as cadastral parcels in forest areas could get the culture code "Parks – Public gardens" as well. To decrease the number of such mistakes, one could apply this rule to urban areas only. For this method a dataset would be needed which contains the borders of urban areas, which is not included in this research project.

Fortunately, the TOP10NL dataset also contains information on the location of parks, as the functional area object class contains the attributes "arboretum" and "park". However, if a park consists of more than one cadastral parcel, only one of these parcels will get the culture code "Parks – Public gardens", as the functional areas from the TOP10NL dataset are point-shaped objects being placed at one location only. As there is no better solution available at the moment, the functional area object class from the TOP10NL dataset is used for the production of the culture code in this research project.

Decision rule 53: If one of the functional area types "arboretum" or "park" from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Parks – Public gardens".

57 Yard – Garden

No direct information for the production of the culture code "Yard – Garden" is offered by the candidate source datasets used in this research project. However, one could use some of the built-up culture codes to produce the culture code. It is assumed that cadastral parcels containing a built-up culture code referring to a residential function, and that also contain surrounding terrain, should get the culture code "Yard – Garden". This means that cadastral parcels containing one of the culture codes "Residential", "Residential (apartment"), "Residential ("agricultural"), "Residential with activity" or "Residential (recreation)", should get the non-built-up culture code "Yard – Garden".

Decision rule 54: If one of the culture codes "Residential", "Residential (apartment"), "Residential ("agricultural"), "Residential with activity" or "Residential (recreation)" is located on a cadastral parcel, and when this parcel also contains the TOP10NL terrain class "other", then this cadastral parcel should get the culture code "Yard – Garden".

61 Land (nature)

For the production of the culture code "Land (nature)" the terrain classes from the TOP10NL dataset can be used. The terrain classes referring to nature are "forest", "heath/moor" and "sand". Other object classes from the TOP10NL dataset (functional, administrative and geographical areas) do offer useful information as well, but these objects are point-shaped. As nature areas are normally covering several cadastral parcels, and as the point-shaped objects are located on one of these parcels only, it is assumed that the terrain classes are more useful for the production of this culture code than the other object classes. Furthermore, the BRP and LGN could offer useful information for the production of the culture code "Land (nature)" as well, but they are not available for this research project. Therefore, they are not included in the decision rules. It is assumed that the terrain classes from the TOP10NL offer the same (and enough) information.

Decision rule 55: If one of the terrain classes "forest", "heath/moor" or "sand" from the TOP10NL dataset is located on a cadastral parcel, then this cadastral parcel should get the culture code "Land (nature)".

62 Land (farming)

For the production of the culture code "Land (farming)" the terrain class "cropland" from the TOP10NL dataset can be used. If this terrain class is located on a cadastral parcel, the parcel should be assigned this culture code.

Decision rule 56: If the terrain class "cropland" from the TOP10NL dataset is located on a cadastral parcel, then this cadastral parcel should get the culture code "Land (farming)".

63 Land (grassland)

The terrain class "grassland" from the TOP10NL dataset can be used for the production of the culture code "Land (grassland)".

Decision rule 57: If the terrain class "grassland" from the TOP10NL dataset is located on a cadastral parcel, then this cadastral parcel should get the culture code "Land (grassland)".

66 Land (cultivation)
The terrain classes “tree nursery” and “fruit farm” can be used for the production of the culture code “Land (cultivation)”.

Decision rule 58: If one of the terrain classes “tree nursery” or “fruit farm” from the TOP10NL dataset is located on a cadastral parcel, then this cadastral parcel should get the culture code “Land (cultivation)”.

76 Culture
Three object classes from the TOP10NL dataset offer information for the production of the culture code “Culture”. These three object classes and the interesting attributes are the following:

1. Water object – function: fish ladder;
2. Cultural landscape object: memorial site/monument and ‘hunebed’;
3. Functional area: burial mound, open air museum and open air theatre.

Decision rule 59: If a cadastral parcel contains a water object with the function “fish ladder” from the TOP10NL dataset, then this cadastral parcel should get the culture code “Culture”.

Decision rule 60: If one of the cultural landscape objects “memorial site/monument” or “hunebed” from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Culture”.

Decision rule 61: If one of the following functional area types from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code “Culture”:

- Burial mound;
- Open air museum;
- Open air theatre.

85 Port
Four object classes from the TOP10NL dataset offer information for the production of the culture code “Port”. These four object classes and the interesting attributes are the following:

1. Water object – function: port;
2. Terrain: landing-stage/pier;
3. Cultural landscape object: landing-stage/pier;
4. Functional area: port, yacht harbor and shipyard.

Decision rule 62: If a cadastral parcel contains a water object with the function “port” from the TOP10NL dataset, then this cadastral parcel should get the culture code “Port”.

Decision rule 63: If a cadastral parcel contains the terrain class “landing-stage/pier” from the TOP10NL dataset, then this cadastral parcel should get the culture code “Port”.

Decision rule 64: If a cadastral parcel contains the cultural landscape object "landing-stage/pier" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Port".

Decision rule 65: If one of the following functional area types from the TOP10NL dataset is located on a cadastral parcel, then this parcel should get the culture code "Port":
- Port;
- Yacht harbor;
- Shipyard.

87 Structures – Water works

Three object classes from the TOP10NL dataset offer information for the production of the culture code "Structures – Water works". These three object classes and the interesting attributes are the following:

1. Water object – physical appearance: in a sluice;
2. Cultural landscape object: pumping station, sluice door, water barrier and dam;
3. Functional area: sluice.

Decision rule 66: If a cadastral parcel contains a water object with the physical appearance "in a sluice" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Structures – Water works".

Decision rule 67: If a cadastral parcel contains one of the cultural landscape objects pumping station, "sluice door", "water barrier" or "dam" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Structures – Water works".

Decision rule 68: If a cadastral parcel contains the functional area "sluice" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Structures – Water works".

89 Water

The TOP10NL object class "water object" offers the information needed to produce the culture code "Water". In the object class a number of attributes are distinguished for the type of water. The following attributes can be used for the production of the culture code: "water stream", "lake, fen, etc.", "ditch" and "sea".

Decision rule 69: If a cadastral parcel contains a water object with the type of water "water stream", "lake, fen, etc.", "ditch" or "sea" from the TOP10NL dataset, then this cadastral parcel should get the culture code "Water".

94 Funeral service

For the production of the culture code "Funeral service" the functional area object class from the TOP10NL dataset can be used. This object class contains the attribute "cemetery". If this

functional area type is located on a cadastral parcel, then this parcel should get the culture code “Funeral service”. Another possible source for the production of this culture code could be the TOP10NL terrain class “Cemetery”. When this terrain class is located on the cadastral parcel, then this cadastral parcel should also get the culture code “Funeral service”. This leads to the next two decision rules:

Decision rule 70: If a cadastral parcel contains the functional area “cemetery” from the TOP10NL dataset, then this cadastral parcel should get the culture code “Funeral service”.

Decision rule 71: If a cadastral parcel contains the terrain class “cemetery” from the TOP10NL dataset, then this cadastral parcel should get the culture code “Funeral service”.

4.2.4 Special properties

The seventy-one decision rules, which are presented in sections 4.2.1, 4.2.2 and 4.2.3, should help to define one or two culture codes for every cadastral parcels. However, it is expected that in some cases the decision rules are not giving enough information for defining a culture code. In those cases, the cadastral parcels should be assigned the culture code “Special properties”. In other words: those cadastral parcels not having a culture code after applying the formulated decision rules should get the culture code “Special properties”. This counts for both built-up and non-built-up culture codes. Therefore, one cadastral parcel should get both the built-up culture code “Special properties” and the non-built-up culture code “Special properties”.

Decision rule 72: If a cadastral parcel has not been given a culture code after applying the decision rules formulated above, then this cadastral parcel should get the culture code “Special properties”.

4.3 – Building a decision tree

In this section, the decision rules formulated in the previous section, are integrated into a decision tree, explained in section 4.3.1. In this decision tree, priority levels are assigned to certain information sources. These priorities are described in this section too (section 4.3.2). Because of the large number of decision rules, the decision tree has a very large size. In order to make the decision tree readable and understandable it has been cut into nine pieces (figures 4.8-4.16).

Figure 4.8 is the first part of the decision tree and shows how the distinction is made between cadastral parcels that should get a only a built-up culture code, only a non-built-up culture code, or both a built-up and a non-built-up code. If more than 90% of a cadastral parcel is covered by a BAG building, then the parcel should only get a built-up culture code. If less than 2% of the parcel is covered by a BAG building, then the parcel should only get a non-built-up culture code. For all percentages in between the parcel should get both a built-up and a non-built-up culture code. The TOP10NL dataset also contains the location and shape of buildings. However, as the BAG dataset is the key register of buildings (and addresses) this dataset is used for the purpose of defining the *bebouwingscode*. As has been explained in section 3, the BAG dataset has a better positional accuracy and is more up-to-date than the TOP10NL dataset.

Figures 4.9-4.12 represent the part of the decision tree for those cadastral parcels which should get a built-up culture code. The highest priority, in this case, is given to the purpose of use for residential objects in the BAG culture code. Again, this dataset is given priority over the TOP10NL, as it is the Dutch key register of buildings (see the previous paragraph). In this case, the BAG dataset is also given priority over the TOP10NL as the object class type of building from the TOP10NL is expected to be far from complete, because the attributes are optional. Figure 4.9 shows how the purposes of use from the BAG dataset (and the apartment rights from the AKR dataset for residential objects with the purpose of use 'residential') should be used to create built-up culture codes. Next, figure 4.10 shows how the object class "type of building" from the TOP10NL dataset can be used to produce built-up culture codes. Figure 4.11 presents the part of the decision tree dealing with the creation of built-up culture codes by making use of the TOP10NL object classes "cultural landscape object" and "functional area". It shows, too, that those parcels still not having a culture code, after using the BAG and TOP10NL dataset, should get the culture code "Special properties". Finally, figure 4.12 shows which cadastral parcels get the culture code "Land (new construction activity)" and which the culture code "Land (new construction residential)" when the status of the residential object is either "building permission granted" or "building process started".

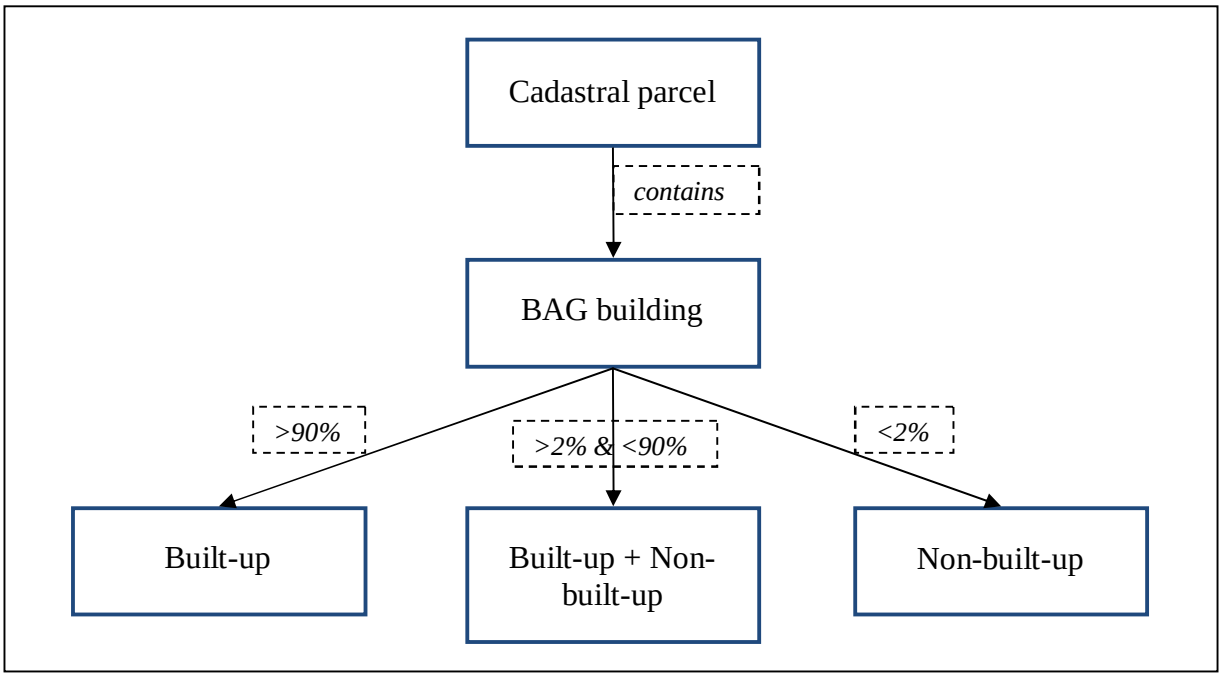


Figure 4.8: Decision tree part I: defining whether the cadastral parcel should get a built-up or non-built-up culture code.

Figures 4.13-4.16 present the part of the decision tree which shows how the non-built-up culture codes can be produced. First, in figure 4.13, for a number of cadastral parcels the non-built-up culture code can be produced by making use of the built-up culture code already available for that parcel.

Second, if none of these built-up culture codes is available, then the TOP10NL object class "functional area" can be used for the production of the non-built-up culture code (figure 4.14). This object class is given a high priority level, as it has a point geometry. As a result, it is not possible to calculate percentages of overlap with the cadastral parcels.

The same counts for the object class "cultural landscape object" (point geometry) and "railroad section" (line geometry). Road sections are also included here, although they have both a line and polygon geometry. With the line geometry it is not possible to calculate

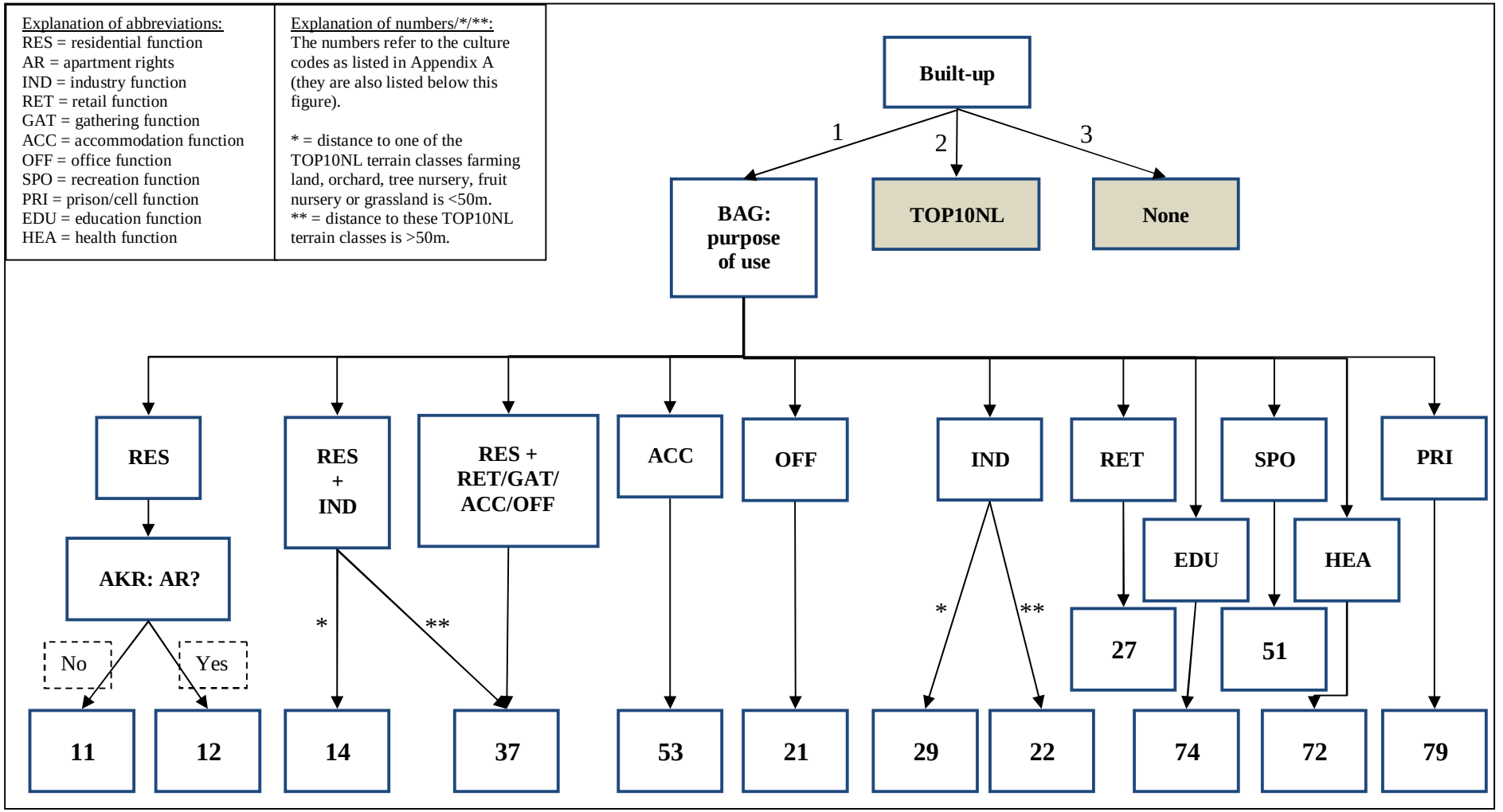
percentage of overlap with parcels, but with polygons it is. However, on larger parcels, roads are expected to have less overlap with the parcels than, for example, grassland. As the roads are sometimes more important, it is decided to give the road sections a higher priority level than the terrain types. Therefore, third, in figure 4.15 it is shown how the TOP10NL object classes “railroad section”, “road section” and “cultural landscape object” can be used for assigning a non-built-up culture code to a cadastral parcel.

Finally, in figure 4.16 the TOP10NL object classes “terrain classes” and “water object” are used for the production of a non-built-up culture code for a parcel. These object classes all have a polygon geometry, which makes it possible to calculate a percentage of overlap with the cadastral parcels. The terrain type (or water) having the highest percentage of overlap is assigned to the cadastral parcel as a culture code.

4.4 – Discussion

By using the decision rules and tree, it is possible to create most of the culture codes listed in Appendix A. Only the culture codes “Pipes – Tubes” and “Activity (glasshouse)” do not appear in the decision tree. For these two culture codes additional information should be used, if it is necessary to keep these culture codes.

However, that only the culture codes “Pipes – Tubes” and “Activity (glasshouse) cannot be produced with the BAG and TOP10NL dataset, does not mean there are no other issues. A serious issue is the completeness of the to be produced culture code dataset. Especially those culture codes completely relying on attribute classes being optional (physical appearance of water objects and type of buildings in the TOP10NL dataset) could prove to be far from complete when applying the decision rules and decision tree to the whole country. Built-up culture codes completely relying on these optional information sources are: Storeroom – shelter (garage-shed), Activity (utility), Activity (catering), Parking, Defense, Public transport, Air traffic, Culture, Police – Fire service, Religion, Funeral service and Port. There are no non-built-up culture codes completely relying on optional information sources (only the culture code “Structures – Water works” is partly relying on optional information).



Culture codes: 11 Residential, 12 Residential (apartment), 14 Residential (agricultural), 21 Activity (office), 22 Activity (industry), 27 Activity (retail trade), 29 Activity (agricultural), 37 Residential with activity, 51 Recreation – Sport, 53 Residential (recreation), 72 Health, 74 Education, 79 Justice.

Figure 4.9: Decision tree part II: defining a built-up culture code by using the purpose of use from the BAG dataset.

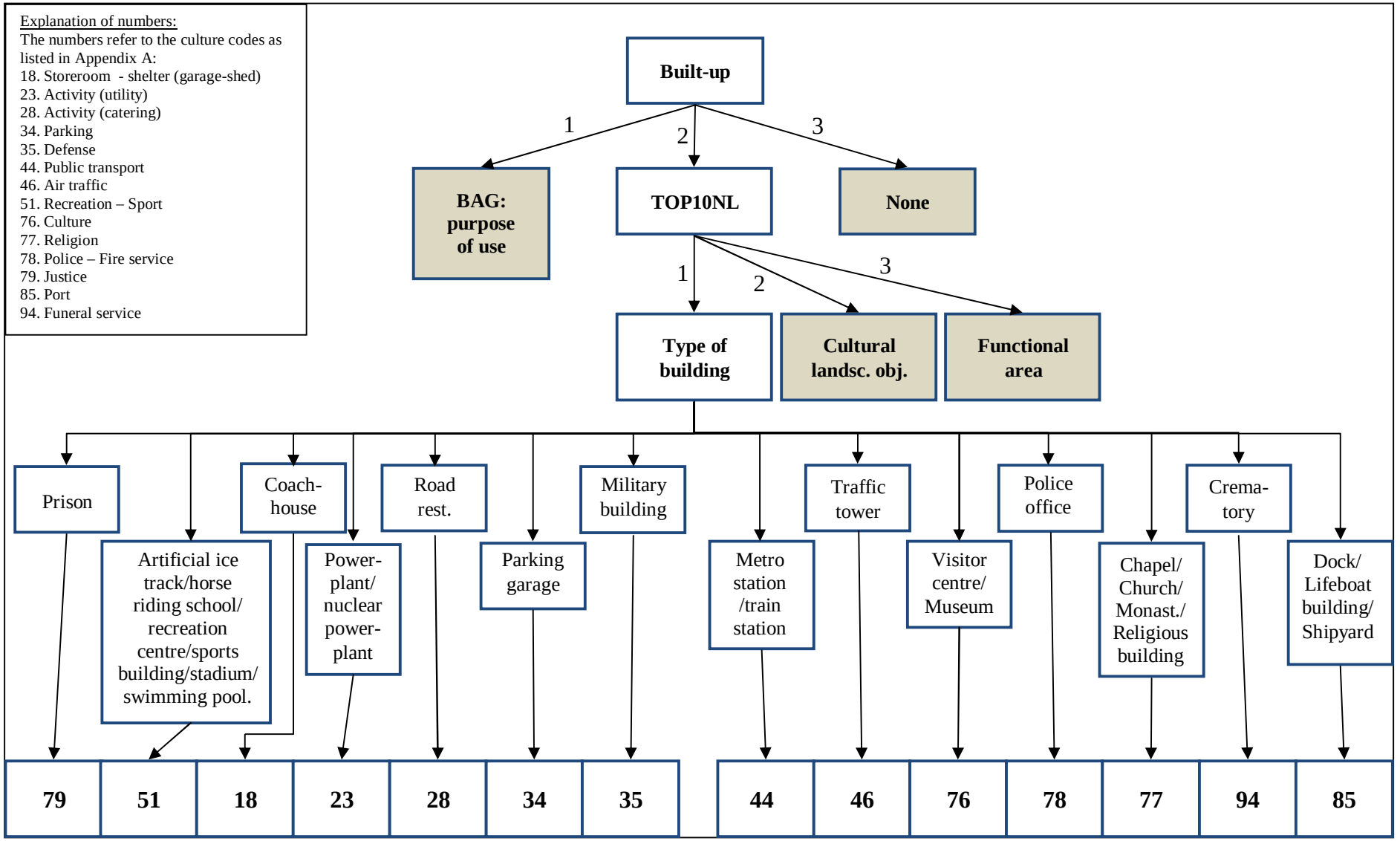


Figure 4.10: Decision tree part III: defining a built-up culture code by making use of the TOP10NL attribute class "type of building".

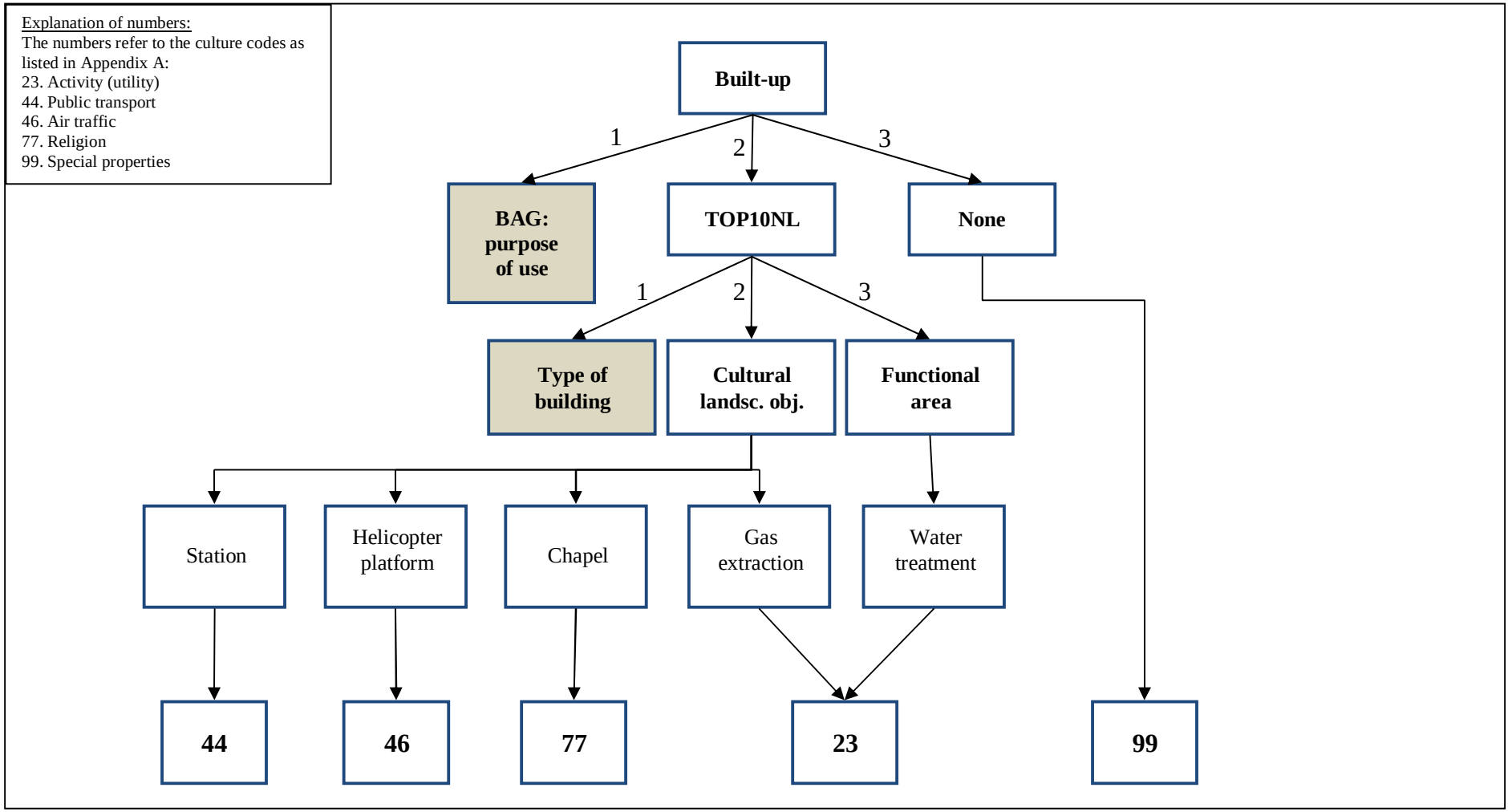


Figure 4.11: Decision tree part IV: defining a built-up culture code by making use of the TOP10NL attribute classes "cultural landscape object" and "functional area".

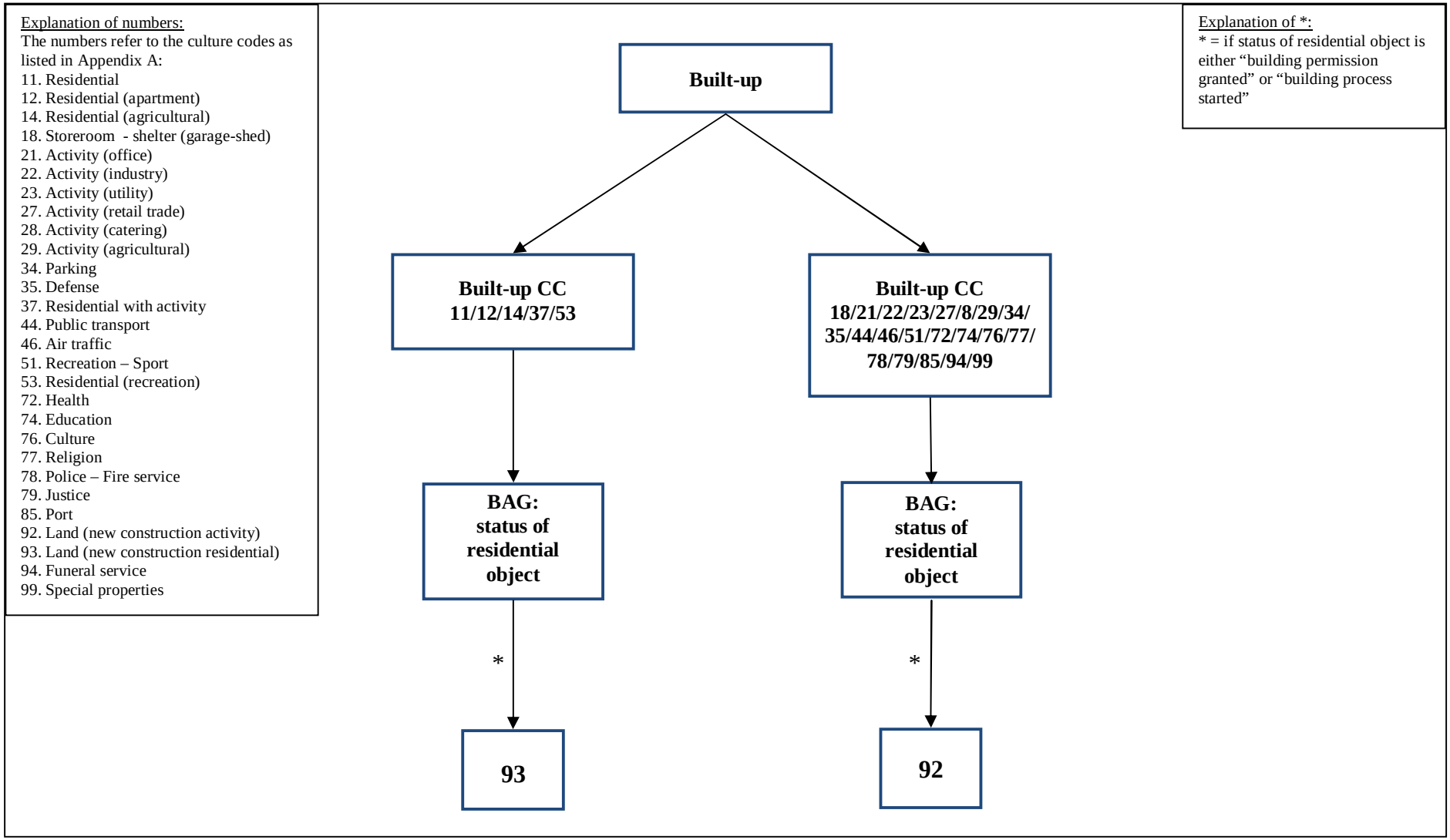


Figure 4.12: Decision tree part V: defining a built-up culture code by making use of the BAG residential object statuses "building permit granted" and "building process started".

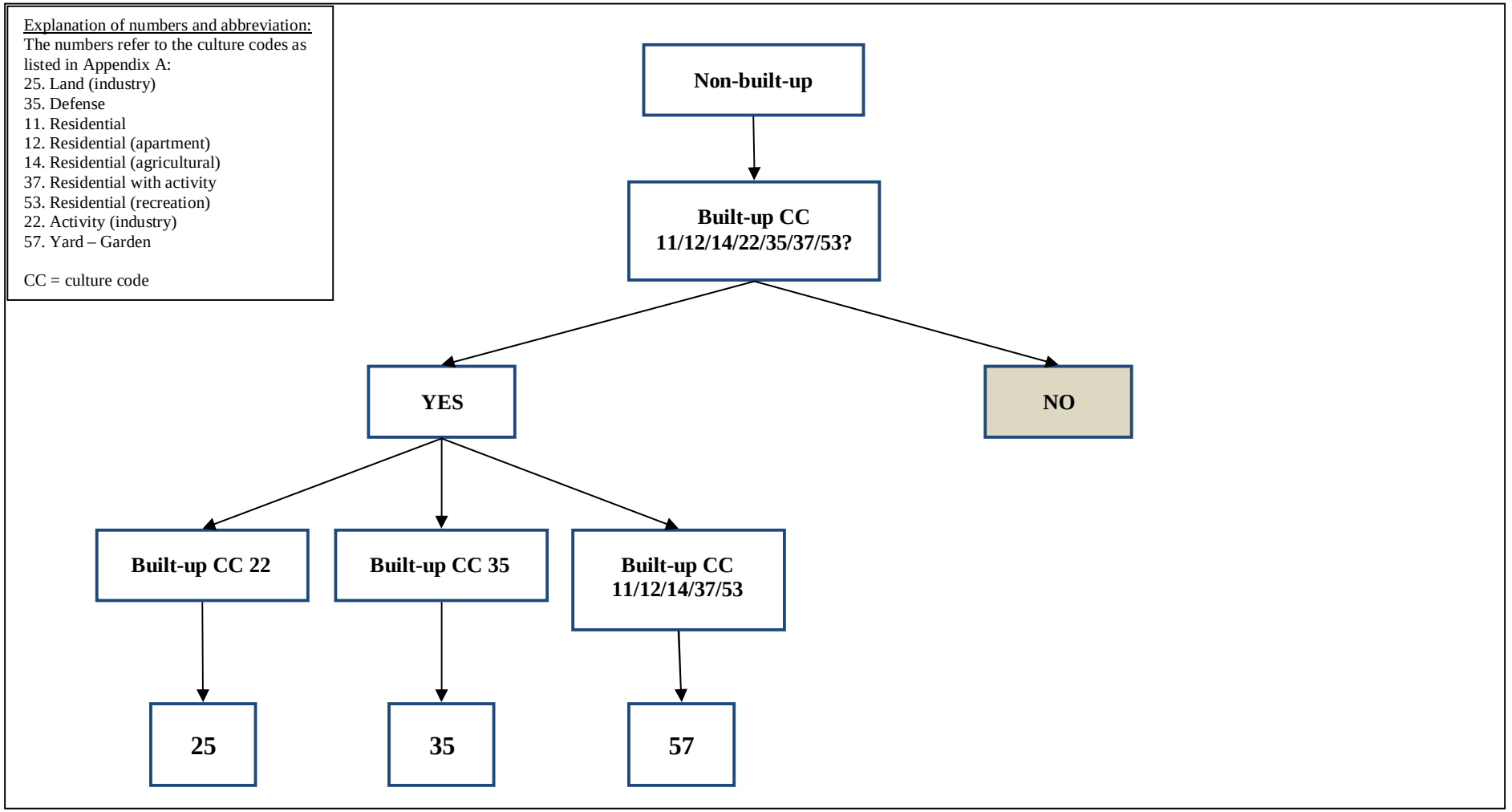


Figure 4.13: Decision tree part VI: defining a non-built-up culture code by making use of the built-up culture codes.

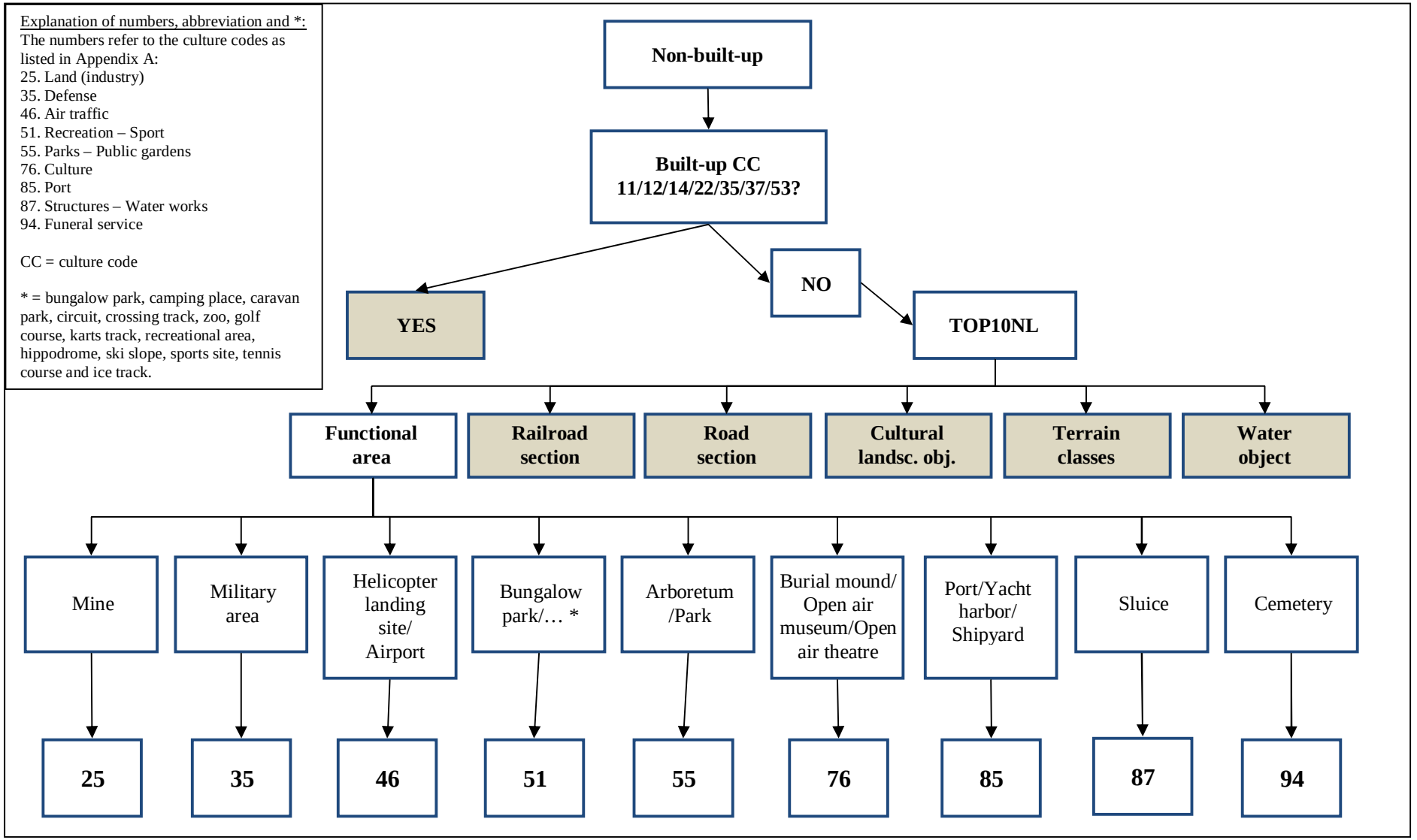


Figure 4.14: Decision tree part VII: defining a non-built-up culture code by making use of the TOP10NL object class "functional area".

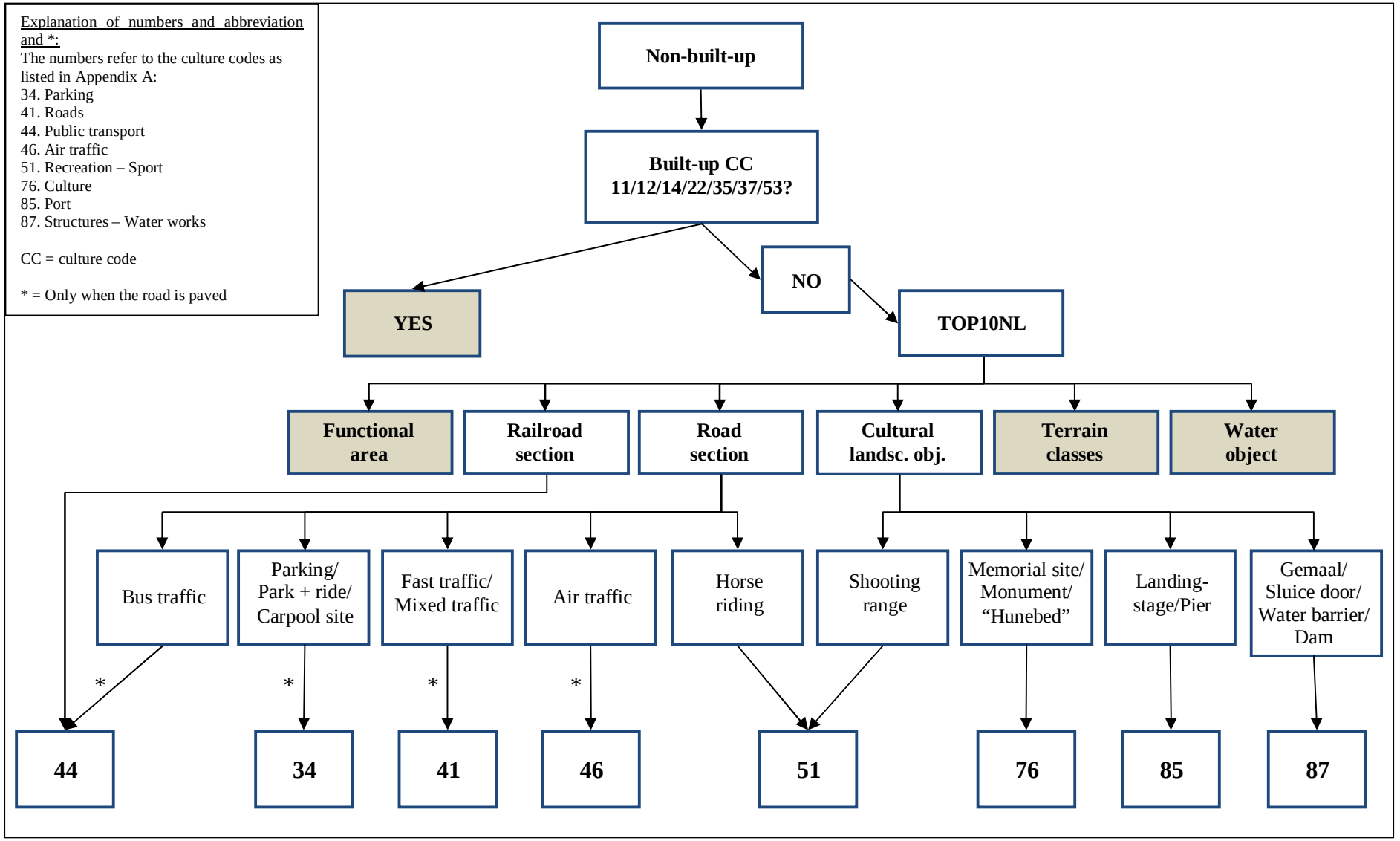


Figure 4.15: Decision tree part VIII: defining a non-built-up culture code by making use of the TOP10NL object classes "railroad section", "road section" and "cultural landscape object".

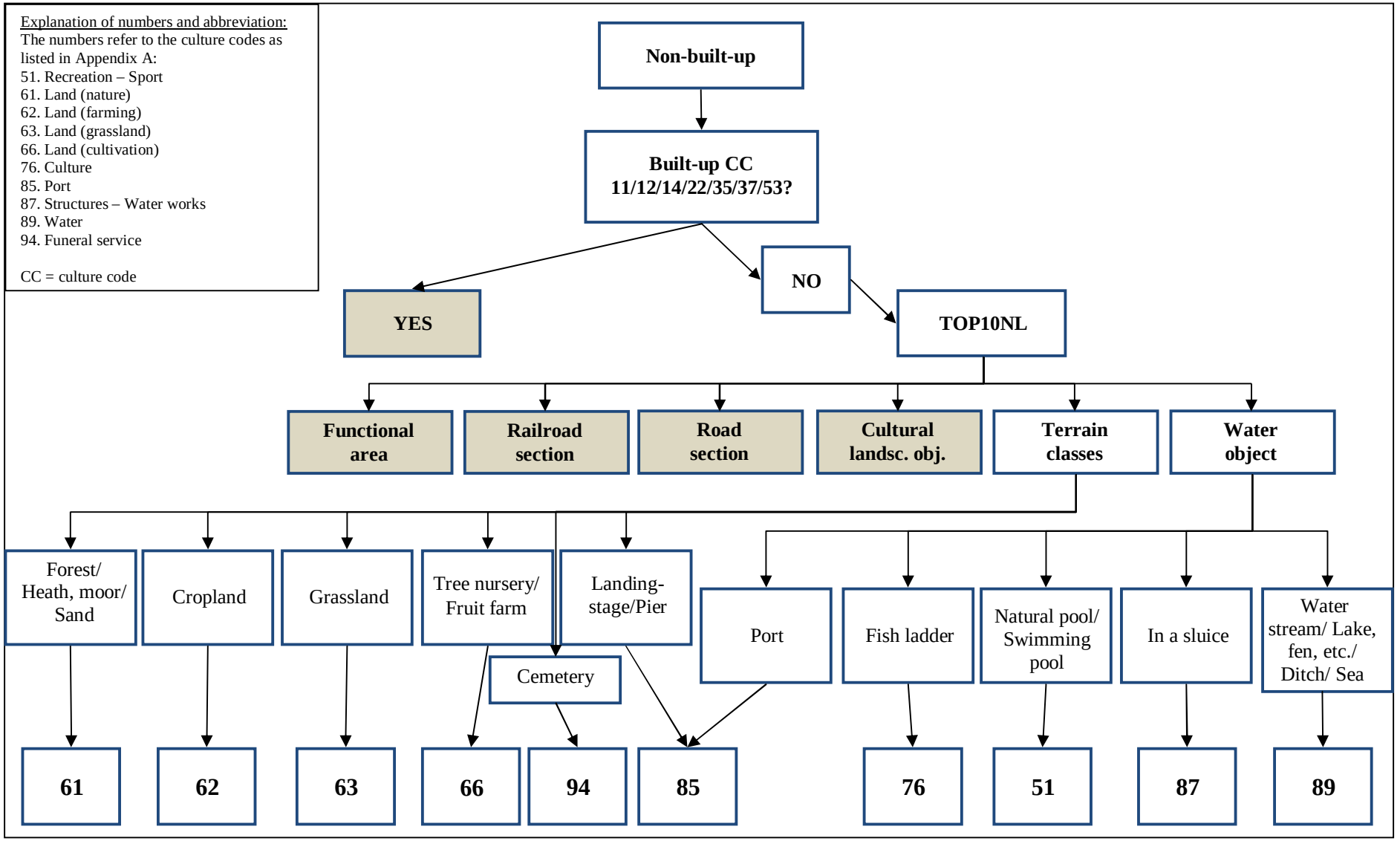


Figure 4.16: Decision tree part IX: defining a non-built-up culture code by making use of the TOP10NL object classes "terrain classes" and "water object".

Chapter 5 – Proof of concept: testing the decision rules

In chapter 4 a large number of decision rules has been formulated. Furthermore, an extensive decision tree was created which integrates these decision rules. The current chapter deals with testing these decision rules and the decision tree, by means of a proof of concept. The original plan has been to perform this proof of concept for a complete municipality (Enschede), but because of serious performance problems – e.g. computer crashing several times after a very long calculation – it has been decided to limit the proof of concept to a number of smaller test areas. In order to get a relatively complete view on the usefulness of the decision rules and decision tree, these test areas should cover the most common situations and should have very different characteristics. Therefore, it has been chosen to perform the proof of concept on a rural area, an area in the city centre, a living area and an industrial/harbor area, all located within the municipality of Enschede (as the spatial data for this complete municipality has been collected). Figure 5.1 shows an overview map of the municipality with the locations of the test areas.

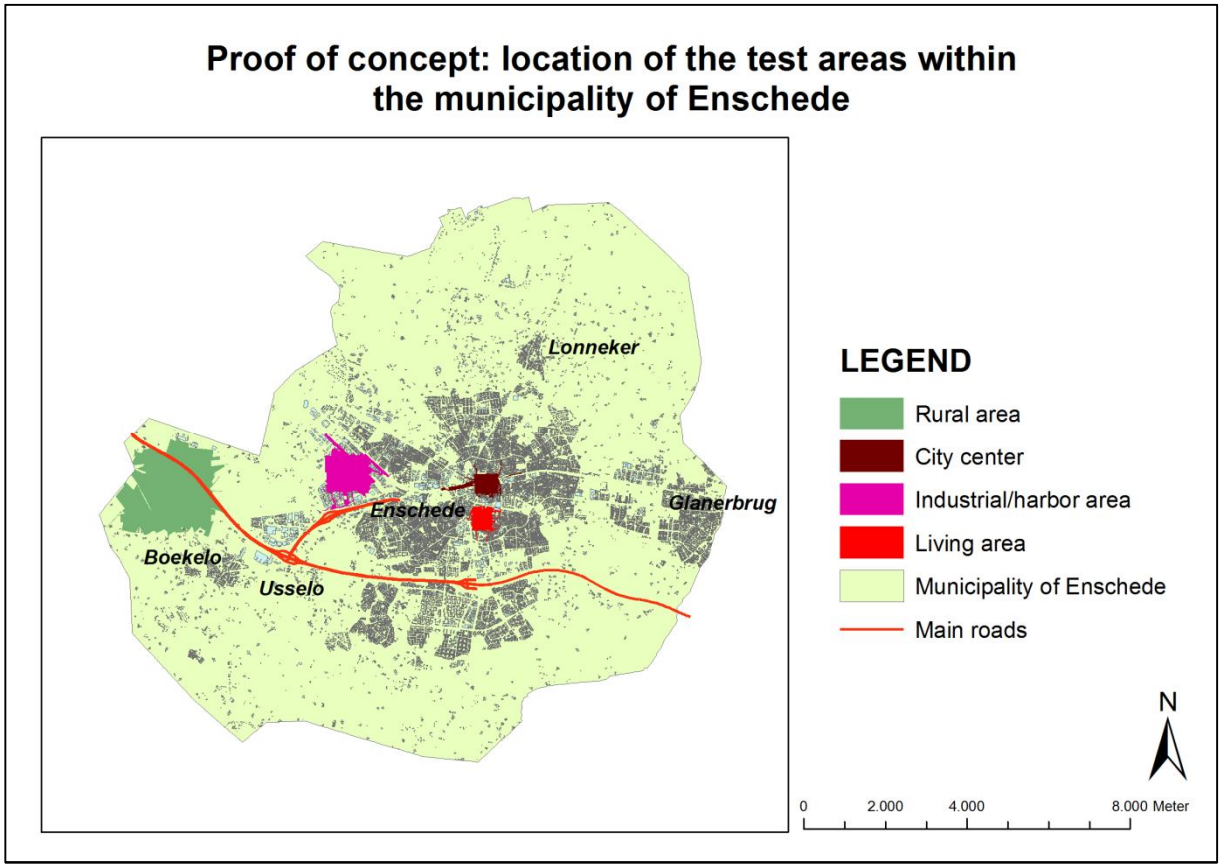


Figure 5.1: Location of the test areas for the proof of concept.

The following sections present the case study and quality analysis on the results for each of these test areas:

- Section 5.1: Rural area
- Section 5.2: City centre
- Section 5.3: Living area
- Section 5.4: Industrial/Harbor area

A detailed description of the method applied for both the case studies and the quality analyses is included in the next sections. However, an overview of the most important steps to be carried out during the proof of concept are explained here.

Within the case study, one can distinguish between three main steps. These steps and their sub-steps are the following:

1. Defining whether a cadastral parcel should get only a built-up, only a non-built-up, or both a built-up and a non-built-up culture code.
    - Calculate the percentage of overlap of BAG buildings with cadastral parcels.
    - Create a layer containing those parcels that should get a built-up culture code (more than 2% overlap) (see the part of the decision tree in figure 4.8).
    - Create a layer containing those parcels that should get a non-built-up culture code (less than 90% overlap) (see figure 4.8).
2. Defining the built-up culture codes for those parcels which should get one.
    - Use the purposes of use in the BAG dataset by joining them to the cadastral parcels and define the culture code (see figure 4.9).
    - Use the TOP10NL object class type of building to define the culture codes for remaining cadastral parcels (see figure 4.10).
    - Use the TOP10NL object classes cultural landscape object and functional area to define the culture codes for remaining cadastral parcels (see figure 4.11).
    - Use the status of the BAG residential objects to find out whether cadastral parcels should get the culture code “Land (new construction activity)” or “Land (new construction residential)” (see figure 4.12).
3. Defining the non-built-up culture codes for those parcels which should get one.
    - Use the built-up culture codes to define non-built-up culture codes (see figure 4.13).
    - Use the TOP10NL object class functional area to define the culture codes for the remaining parcels (see figure 4.14).
    - Use the TOP10NL object class railroad section to define culture codes (figure 4.15).
    - Use the TOP10NL object class road section to define culture codes (figure 4.15).
    - Use the TOP10NL object class cultural landscape objects to define culture codes (figure 4.15).
    - Use the TOP10NL terrain classes and water section to calculate the percentages of overlap with the cadastral parcels and use these percentages to define the culture codes for the remaining parcels (figure 4.16). The terrain type with the highest overlap percentage is used for giving a culture code to this parcel.

For the quality analysis a number of steps can be distinguished as well:

1. For every built-up and non-built-up culture code present in the case study area a number of cadastral parcels has to be selected.
2. For these selected it is checked whether the calculated culture codes give the same land use type as in the real-world situation.
3. For the culture codes showing the worst results it is analyzed why they are showing bad results.

The steps for the case study and the quality analysis, as described above, are used in all case study areas. The precise methods applied and the results are presented in the next sections (5.1-5.4).

5.1 – POC: rural area

In this section, the results for the rural area are presented (section 5.1.1), as well as a quality analysis on the results (section 5.1.2). For both the case study and quality analysis, a detailed description of the method applied is included.

5.1.1 Case study

Built-up, non-built-up or combined

The first step is to calculate the area of the parcels (in square meters). This is done by adding a new field to the attribute table of the parcel layer and use the "Calculate geometry" option. The result is a new column in the attribute table with the area per parcel.

Second, one needs to define the percentage of the cadastral parcel covered by BAG buildings[13]. For this calculation, the Hawth's Analysis tool "Polygon in Polygon Analysis" is used. This analysis tool calculates the area of the cadastral parcel covered by the BAG building polygons (see figure 5.2 for a screenshot of the tool's window).

When performing this tool, a first problem was encountered. In a number of cases the resulting area was larger than the area of the cadastral parcel, because every BAG building polygon existed three times in the layer. Therefore, the redundant polygons had to be deleted first. After deleting these polygons the Polygon in Polygon Analysis was performed a second time, resulting in the correct area of the cadastral parcels covered by BAG buildings. Next, a new column "Perc_BAG" has been to the attribute table of the cadastral parcels layer. For this new column the "Field Calculator" was used to calculate the percentage of the parcel cover by BAG buildings. For this calculations the following formula was used:

*Perc_BAG = (PIPA_AR / AREA) * 100*

Where:

- Perc_BAG is the percentage of the parcel covered by BAG buildings
- PIPA_AR is the area of the parcel covered by BAG buildings
- AREA is the total area of the parcel

With the calculated percentage it is possible to define the "*bebouwingscode*" (in a new column BEBCODE), which can have the following values, according to the threshold values presented in the previous chapter:

1. Only non-built-up
2. Built-up and non-built-up
3. Only built-up

In this case study the highest percentage of a parcel that is covered by BAG buildings is 67%. This means none of the cadastral parcels has only built-up elements. 234 Cadastral parcels have less than 2% covered by BAG buildings and get only a non-built-up culture code. The remaining 22 parcels should get both a built-up and a non-built-up culture code. Figure 5.3 presents a map showing the distribution of these cadastral parcels.

[13] Buildings which have been demolished already should be removed from the BAG buildings dataset, as they are still present in the dataset. In the case of the rural area, this meant three building features had to be removed from the dataset.

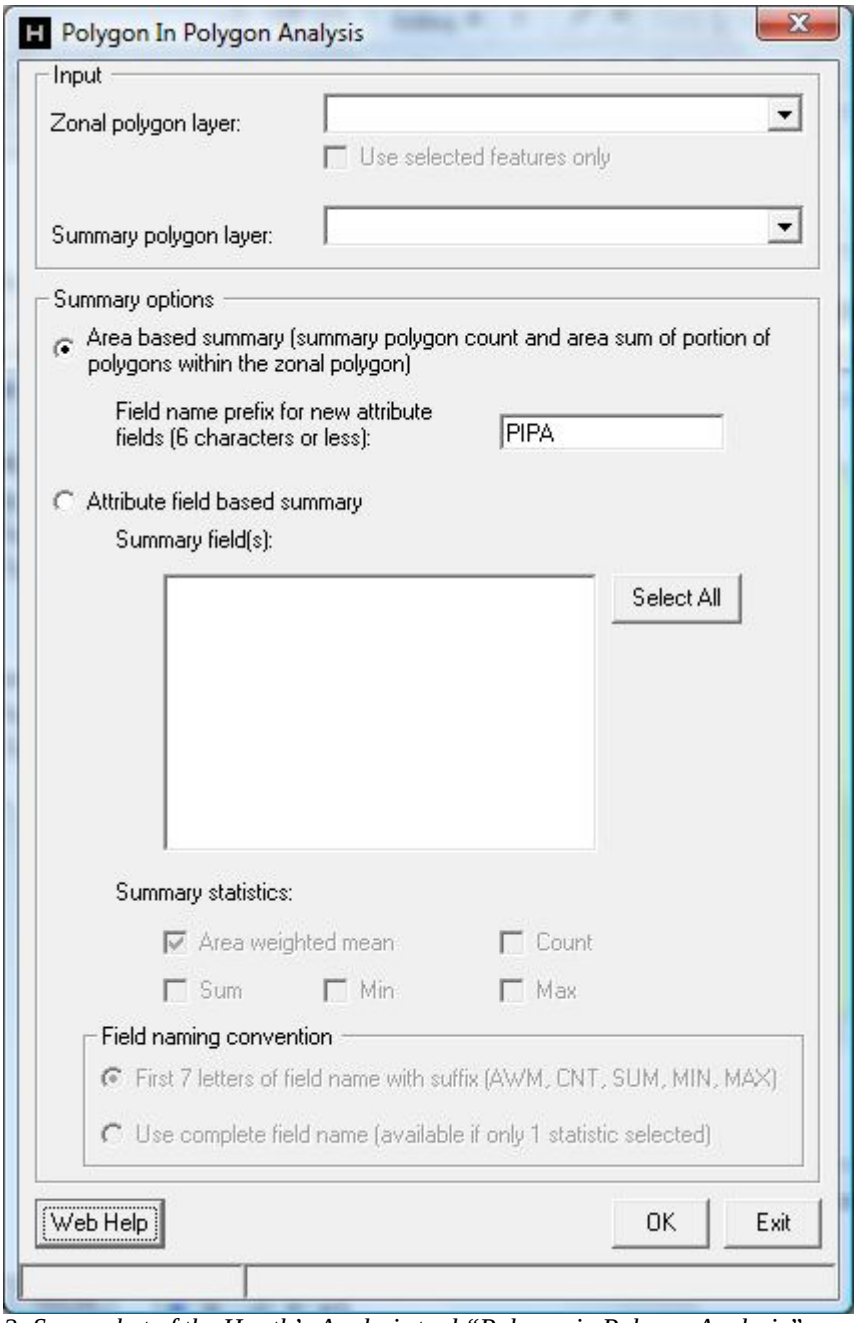


Figure 5.2: Screenshot of the Hawth's Analysis tool "Polygon in Polygon Analysis".

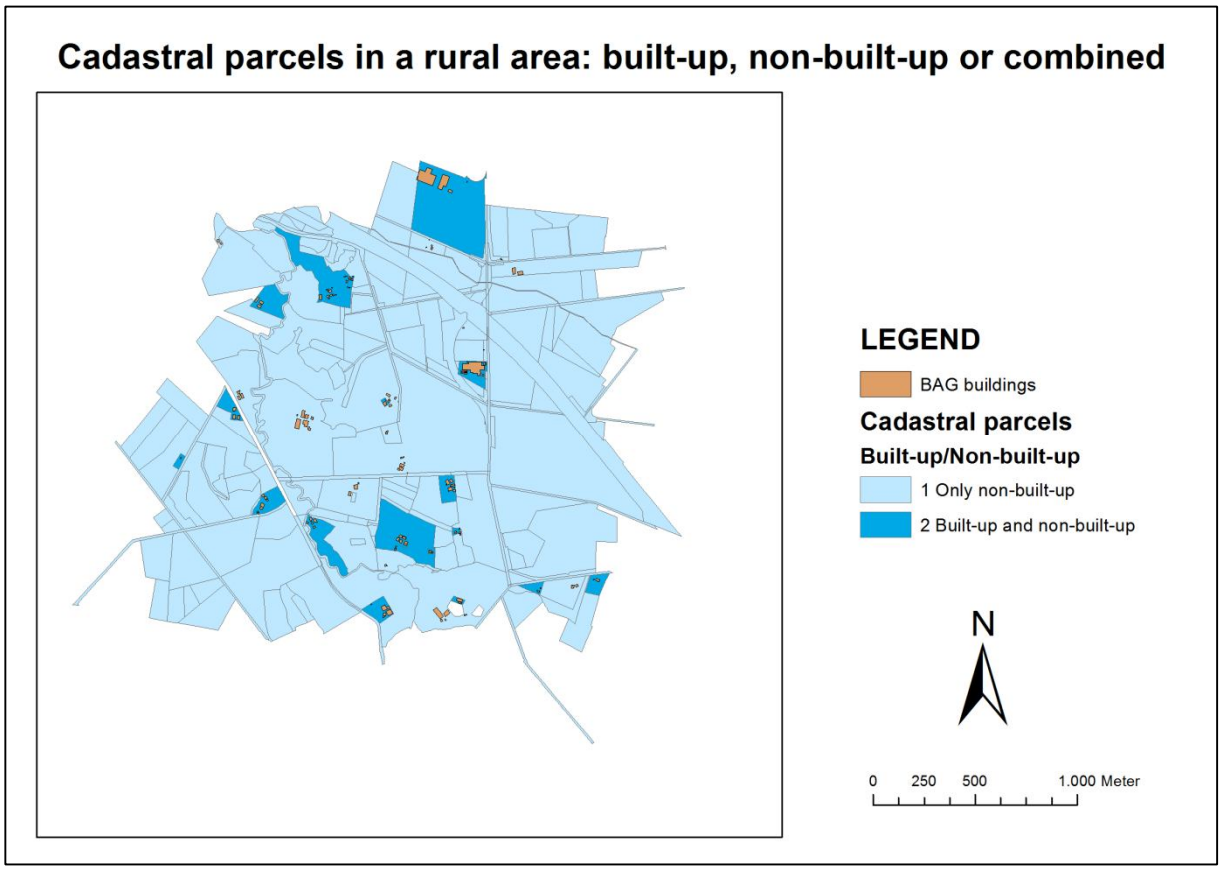


Figure 5.3: Cadastral parcels in a rural area: built-up, non-built-up or combined.

Attributes of Cadastral parcels - built-up

FID	Shape *	Join_Count	Id	Area	PIPA_AR	PIPA_CNT	Perc_BAG	BEBCODE	PNTPOLYCNT	TYPE
2	Polygon	1	0	4048,1399	99,818032	2	2,46578	2	1	Verblijfso
3	Polygon	1	0	7818,6699	203,831682	2	2,60699	2	1	Verblijfso
4	Polygon	1	0	897,76202	176,506875	2	19,660801	2	1	Verblijfso
7	Polygon	1	0	991,172	163,991894	6	16,5453	2	1	Verblijfso
8	Polygon	1	0	21363,801	651,581658	5	3,04993	2	1	Verblijfso
11	Polygon	2	0	12419,7	898,020466	7	7,23061	2	2	Verblijfso
15	Polygon	1	0	1005,94	371,145805	3	36,895401	2	1	Verblijfso
16	Polygon	1	0	8376,21	710,7931	3	8,48586	2	1	Verblijfso
17	Polygon	1	0	639,52502	92,07849	1	14,3979	2	1	Verblijfso
1	Polygon	2	0	10533,3	1453,160594	6	13,7959	2	2	Verblijfso
9	Polygon	2	0	59283,398	1233,017011	11	2,07987	2	2	Verblijfso
13	Polygon	2	0	9847,8799	1141,723028	9	11,5936	2	2	Verblijfso
20	Polygon	3	0	17749,801	668,353327	3	3,76541	2	3	Verblijfso
22	Polygon	3	0	46851,301	1270,042666	17	2,71079	2	3	Verblijfso
19	Polygon	2	0	12666,4	5004,01512	7	39,506199	2	2	Verblijfso
0	Polygon	0	0	1576	459,639905	2	29,165001	2	0	
5	Polygon	0	0	788,17999	109,938349	3	13,9484	2	0	
6	Polygon	0	0	219,66	46,212048	2	21,038	2	0	
14	Polygon	0	0	2180,1001	44,59709	1	2,04564	2	0	
18	Polygon	0	0	500,75699	25,95842	2	5,18384	2	0	
21	Polygon	0	0	14,382	2,692921	1	18,724199	2	0	
23	Polygon	0	0	123073	7374,541461	5	5,99201	2	0	

Record: 1 Show: All Selected Records (0 out of 22 Selected) Options

Figure 5.4: Extract of the attribute table with the column PNTPOLYCNT containing information on the number of residential objects located on a single cadastral parcel.

Built-up culture codes

After defining which cadastral parcels should get a built-up culture code, a new layer is created containing those parcels. The first step, as has been shown in the decision tree, is to use the purpose of use for the residential objects in the BAG dataset. The residential objects are added to the map and the Hawth's Analysis tool "Count Points in Polygons" is used to calculate how many purposes of use are located on the cadastral parcels which should get a built-up culture code. Figure 5.4 shows the result of this calculations, with the column "PNTPOLYCNT" containing the number of residential objects located on a parcel.

The next step is a spatial join ("one-to-one") between the cadastral parcels and the residential objects. The result is shown in figure 5.5. The purpose of use is added to the attribute table in the column "GEBRUIKSDO". However, if more than one residential object is located on the cadastral parcel, then only one of these is joined to the parcel, meaning that only one purpose of use is shown in the table. For those cadastral parcels only containing one residential object one can simply use the purpose of use in the attribute table to define the culture code. For those parcels containing more than one residential object one has to have a look at the other residential objects too. In case a combination of the purposes of use "residential function" and "retail function"/"gathering function"/"accommodation function"/"office function" is found at a parcel, then the parcel should get the culture code "Residential with activity". In case a combination of the purposes of use "residential function" and "industrial function" is found at a cadastral parcel, then the terrain classes from the TOP10NL dataset have to be used for determining the culture code. If the residential objects are within a 50m distance from an agricultural land use type (farming land, orchard, tree nursery, fruit nursery or grassland), then the parcels gets the culture code "Residential (agricultural)". Otherwise it gets the culture code "Residential with activity)". In the rural area this combination was always within a 50m distance of an agricultural land use type. As a result, these parcels are given the culture code "Residential (agricultural)".

Attributes of Cadastral parcels - built-up

FID	Shape *	Join_Count	Id	Area	PIPA_AR	PIPA_CNT	Perc_BAG	BEBCODE	PNTPOLYCNT	TYPE_OBJEC	STATUS	GEBRUIKSDO
2	Polygon	1	0	4048,1399	99,818032	2	2,46578	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
3	Polygon	1	0	7818,6699	203,831682	2	2,60699	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
4	Polygon	1	0	897,76202	176,506875	2	19,660801	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
7	Polygon	1	0	991,172	163,991894	6	16,5453	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
8	Polygon	1	0	21363,801	651,581658	5	3,04993	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
11	Polygon	2	0	12419,7	898,020466	7	7,23061	2	2	Verblijfsobject	Pand in gebruik	woonfunctie
15	Polygon	1	0	1005,94	371,145805	3	36,895401	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
16	Polygon	1	0	8376,21	710,7931	3	8,48586	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
17	Polygon	1	0	639,52502	92,07849	1	14,3979	2	1	Verblijfsobject	Pand in gebruik	woonfunctie
1	Polygon	2	0	10533,3	1453,160594	6	13,7959	2	2	Verblijfsobject	Pand in gebruik	woonfunctie
9	Polygon	2	0	59283,398	1233,017011	11	2,07987	2	2	Verblijfsobject	Pand in gebruik	woonfunctie
13	Polygon	2	0	9847,8799	1141,723028	9	11,5936	2	2	Verblijfsobject	Pand in gebruik	woonfunctie
20	Polygon	3	0	17749,801	668,353327	3	3,76541	2	3	Verblijfsobject	Pand in gebruik	woonfunctie
22	Polygon	3	0	46851,301	1270,042666	17	2,71079	2	3	Verblijfsobject	Pand in gebruik	industriefunctie
19	Polygon	2	0	12666,4	5004,01512	7	39,506199	2	2	Verblijfsobject	Pand in gebruik	winkelfunctie
0	Polygon	0	0	1576	459,639905	2	29,165001	2	0			
5	Polygon	0	0	788,17999	109,938349	3	13,9484	2	0			
6	Polygon	0	0	219,66	46,212048	2	21,038	2	0			
14	Polygon	0	0	2180,1001	44,59709	1	2,04564	2	0			
18	Polygon	0	0	500,75699	25,95842	2	5,18384	2	0			
21	Polygon	0	0	14,382	2,692921	1	18,724199	2	0			
23	Polygon	0	0	123073	7374,541461	5	5,99201	2	0			

Record: 1 Show: All Selected Records (0 out of 22 Selected) Options

Figure 5.5: The attribute table after a spatial join between the cadastral parcels and the residential objects.

During this procedure, nine cadastral parcels in the rural test area have been given the culture code "Residential". As becomes clear from the decision tree one should also take into account whether apartment rights are located on the cadastral parcel. Those parcels should get the culture code "Residential (apartment)" instead of "Residential". Therefore, a dataset with the location of such apartment rights (as point objects) is added to ArcGIS. The parcels with the culture code "Residential" are selected and from this selection, by making use of the "Select by location" tool in ArcGIS, those parcels with apartment rights located on them are selected.

In the rural test area, none of the cadastral parcels with the culture code “Residential” contain apartment rights.

For those cadastral parcels containing no residential objects or only residential objects with the purpose of use “other function” the TOP10NL dataset is used to derive abuilt-up culture code. In the rural area none of the TOP10NL object classes (Type of building, Cultural landscape object and Functional area) contains information which can be used for the production of the Kadaster built-up culture codes for the remaining parcels. As a result, those cadastral parcels requiring a built-up culture code and not yet having one are given the culture code “Special properties”.

A final step is to find out whether there are cadastral parcels containing BAG buildings with either the status “building process started” or the status “building permit granted”. If one of these statuses is found on a cadastral parcel, then the parcel should get either the culture code “Land (new construction activity)” or the culture code “Land (new construction residential)”. In the rural area none of the cadastral parcels contained one of these two statuses.

Figure 5.6 presents a map showing the cadastral parcels having a built-up culture code. It also shows which built-up culture code has been assigned to these parcels.

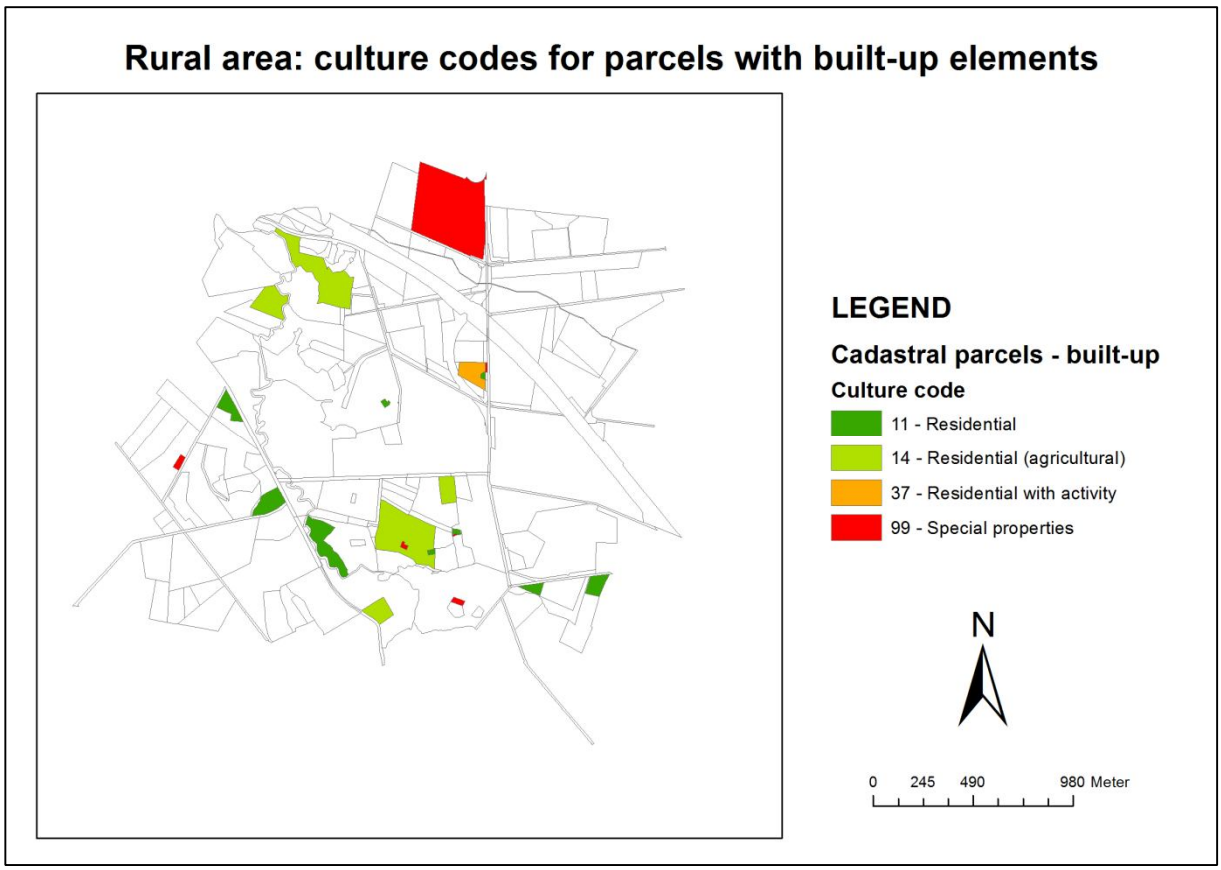


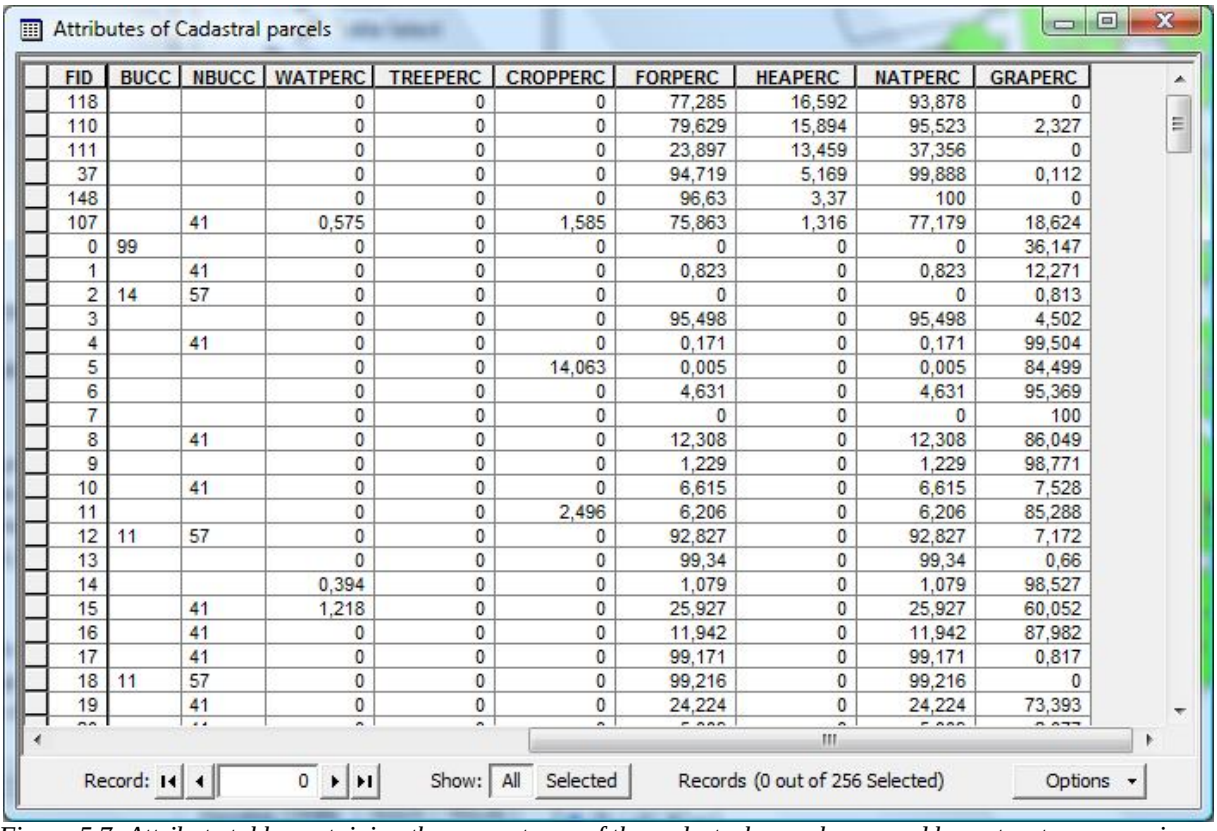
Figure 5.6: Built-up culture codes in the rural area.

Non-built-up culture codes

After defining the built-up culture codes, the non-built-up culture codes have to be defined. As has been argued above, in the rural area all parcels should get a non-built-up culture code. Looking at the decision tree, the first step should be to assign non-built-up culture codes to cadastral parcels containing one of the following built-up culture codes:

- Activity industry → Land (industry)
- Defense → Defense
- Residential/Residential (apartment)/Residential (agricultural)/Residential with activity/Residential (recreational) → Yard – Garden

In the rural area, one can find 15 cadastral parcels containing the culture codes “Residential”, “Residential” (agricultural)” or “Residential with activity”. These parcels are given the non-built-up culture code “Yard – Garden”.



FID	BUCC	NBUCC	WATPERC	TREEPERC	CROPPERC	FORPERC	HEAPERC	NATPERC	GRAPERC
118			0	0	0	77,285	16,592	93,878	0
110			0	0	0	79,629	15,894	95,523	2,327
111			0	0	0	23,897	13,459	37,356	0
37			0	0	0	94,719	5,169	99,888	0,112
148			0	0	0	96,63	3,37	100	0
107		41	0,575	0	1,585	75,863	1,316	77,179	18,624
0	99		0	0	0	0	0	0	36,147
1		41	0	0	0	0,823	0	0,823	12,271
2	14	57	0	0	0	0	0	0	0,813
3			0	0	0	95,498	0	95,498	4,502
4		41	0	0	0	0,171	0	0,171	99,504
5			0	0	14,063	0,005	0	0,005	84,499
6			0	0	0	4,631	0	4,631	95,369
7			0	0	0	0	0	0	100
8		41	0	0	0	12,308	0	12,308	86,049
9			0	0	0	1,229	0	1,229	98,771
10		41	0	0	0	6,615	0	6,615	7,528
11			0	0	2,496	6,206	0	6,206	85,288
12	11	57	0	0	0	92,827	0	92,827	7,172
13			0	0	0	99,34	0	99,34	0,66
14			0,394	0	0	1,079	0	1,079	98,527
15		41	1,218	0	0	25,927	0	25,927	60,052
16		41	0	0	0	11,942	0	11,942	87,982
17		41	0	0	0	99,171	0	99,171	0,817
18	11	57	0	0	0	99,216	0	99,216	0
19		41	0	0	0	24,224	0	24,224	73,393

Figure 5.7: Attribute table containing the percentages of the cadastral parcels covered by water, tree nurseries, cropland, nature (combination of forest and heath) and grassland.

For the remaining cadastral parcels requiring a non-built-up culture code, the TOP10NL dataset is used. The following object classes are used to define the non-built-up culture codes:

- Functional area. No useful information from this object class is found in the rural area.
- Railroad section. No railroad sections are located in the rural area.
- Road section. In the rural area paved road sections with the main use types “fast traffic” and “mixed traffic” are found. First, “Select by attributes” is used to select the parcels which do not yet have a non-built-up culture code. Second, “Select by location” is used to select (from the already selected parcels) the parcels that contain a road section, by making use of the option “are crossed by the outline of”. 80 Cadastral parcels are selected and should get the non-built-up culture code “Roads”, independent of the percentage of overlap.
- Cultural landscape object. Not present in the rural area.
- Terrain classes and water objects. For these object classes the Hawth’s Analysis tool “Polygon in Polygon Analysis” is used to calculate the area of the cadastral parcel covered by these object classes. Afterwards, this area is converted into a percentage of

the parcel. The result is shown in figure 5.7. After calculating these percentages, they are used to define which land use type covers the largest part of the parcel (see figure 5.8 for the selection method). Then the non-built-up culture code belonging to this land use type is assigned to the parcel:

- 89 Water: 0 parcels
- 66 Land (cultivation): 1 parcel
- 62 Land (farming): 7 parcels
- 61 Land (nature): 66 parcels
- 63 Land (grassland): 83 parcels

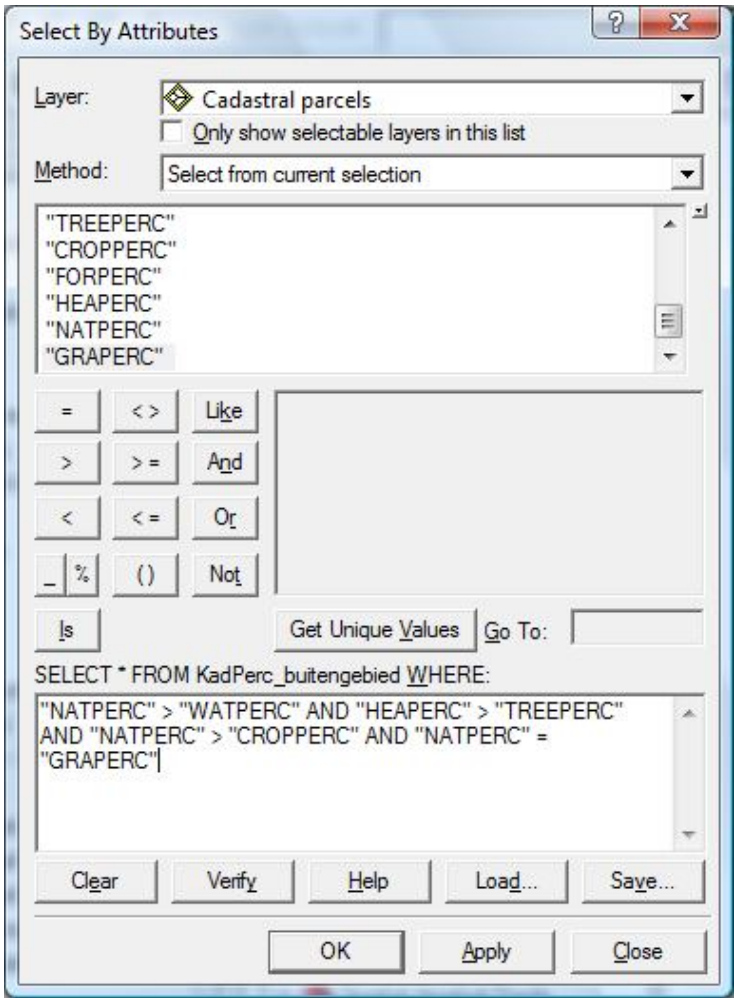


Figure 5.8: Selection method used to select the land use type covering the largest part of the parcel.

The remaining cadastral parcels which should still get a non-built-up culture code, get the culture code “Special properties”. In the rural area, this culture code is assigned to four cadastral parcels.

Figure 5.9 presents a map showing the distribution of the non-built-up culture codes in the rural area. What is directly visible in the map, is the large amount of cadastral parcels being given the culture code “Roads”. The reason for this is, partly, that polygon features of the road sections have been used, which are in some cases just crossing the boundary of a cadastral parcel. In order to decrease the influence of the road sections overlapping the cadastral parcels for only a very small part, it has been chosen to repeat the procedure for defining the non-built-up culture codes, but this time making use of the road section central

lines (and the “Select by location” option “intersect” instead of “are crossed by the outline of”). The results differ quite a lot:

- Roads: 42 parcels
- Water: 1 parcel
- Land (cultivation): 2 parcels
- Land (farming): 8 parcels
- Land (nature): 79 parcels
- Land (grassland): 105 parcels
- Special properties: 4 parcels

Figure 5.10 presents a new map showing the distribution of the non-built-up culture codes.

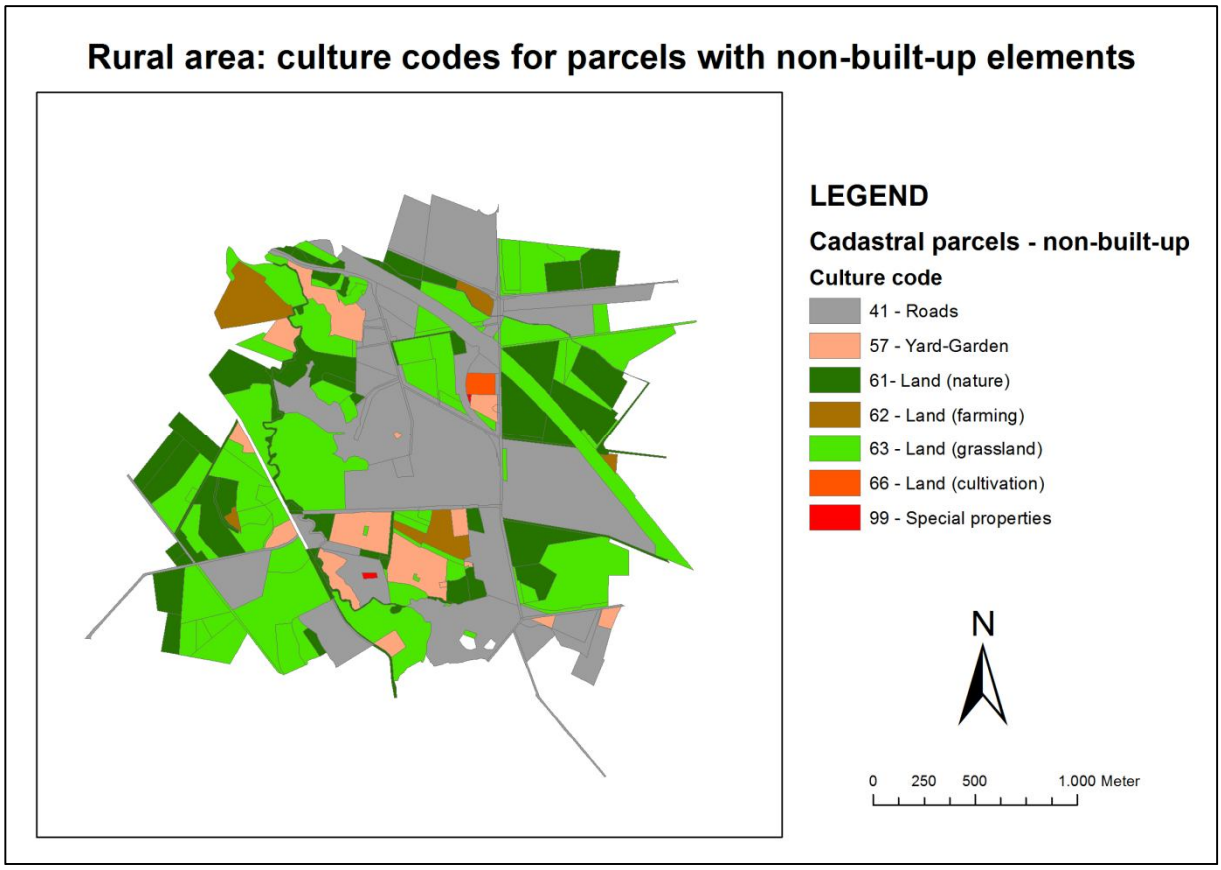


Figure 5.9: Non-built-up culture codes in the rural area.

To sum up, table 5.1 presents the final results for the built-up and non-built-up culture codes in the rural area. It shows for every culture code the number of cadastral parcels it has been assigned to and the total number of built-up and non-built-up culture codes.

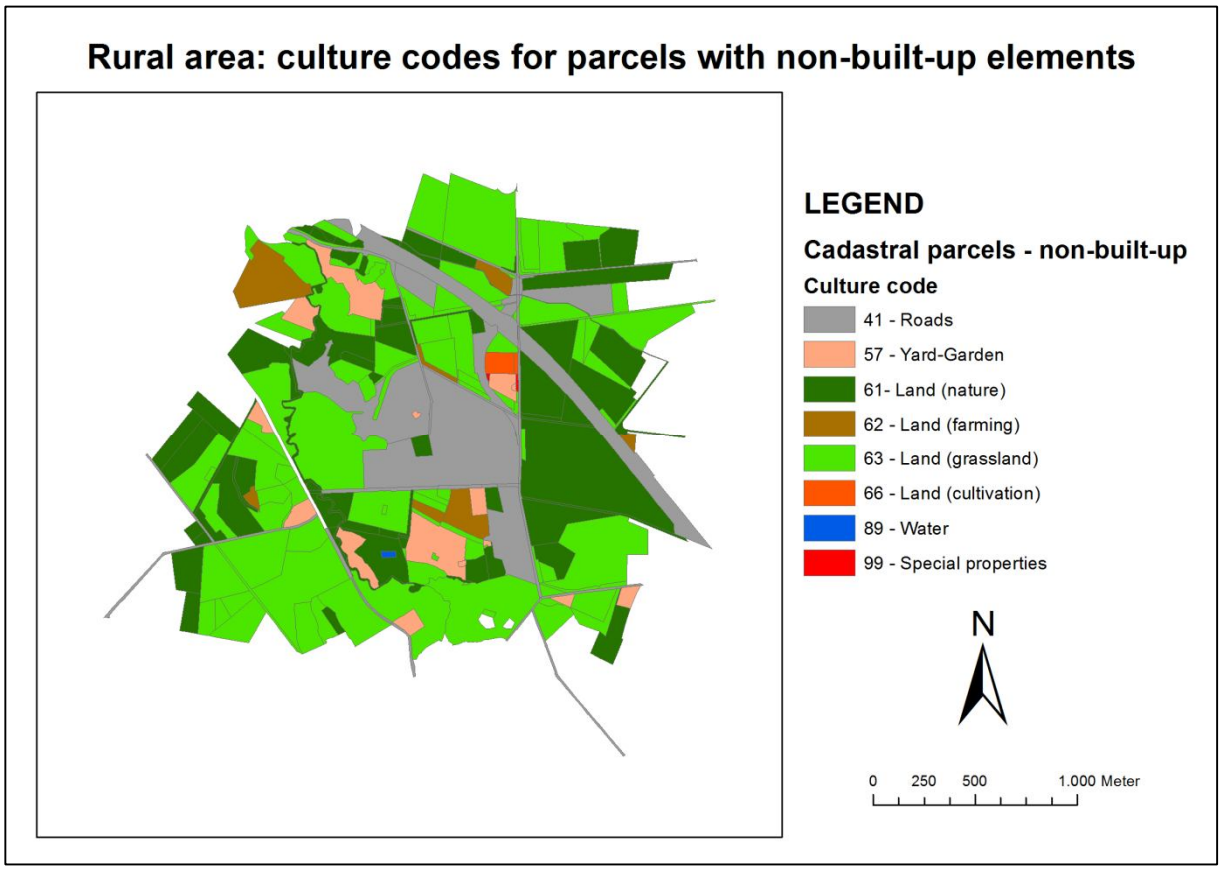


Figure 5.10: Non-built-up culture codes in the rural area when using the roads central lines.

Culture code	Built-up	Non-built-up
11 Residential	9	
14 Residential (agricultural)	5	
37 Residential with activity	1	
41 Roads		42
57 Yard – Garden		15
61 Land (nature)		79
62 Land (farming)		8
63 Land (grassland)		105
66 Land (cultivation)		2
89 Water		1
99 Special properties	7	4
TOTAL	***22***	***256***

Table 5.1: Results for built-up and non-built-up culture codes in rural area.

5.1.2 Quality analysis

Now that the decision rules and decision tree have been tested on the cadastral parcels in a rural test area, the results of this test are validated. In other words, a quality analysis is performed on the results of the production method proposed in this thesis.

This quality analysis is performed by comparing the culture codes found in the previous section (5.1.1) with the real-life situation. The author has visited the test area in order to know what is situated on the cadastral parcel in real-life. Because of time constraints it is not possible to visit all cadastral parcels in the test areas and, therefore, a number of cadastral parcels is selected – for the built-up and non-built-up culture codes – and the results for these

cadastral parcels are validated. A number of rules for the selection of these cadastral parcels have been applied:

1. For every culture code present in the test area one or more cadastral parcels are selected;
2. The number of selected parcels depends on the total number of cadastral parcels which were given this culture code. In general, the following rule is applied: the higher the number of cadastral parcels having a specific culture code, the higher the number of selected parcels for the quality analysis;
3. The selected cadastral parcels have to be directly next to or, at least, visible from public roads, as the author should not risk to be on private property.

Results for built-up culture codes

Figure 5.11 shows the cadastral parcels in the rural test area which have been selected for the quality analysis on the results for the built-up culture codes. The legend contains, next to the explanation of the colors used in the map, the number of selected cadastral parcels per culture code. For example, four cadastral parcels have been selected which are given the culture code "Residential".

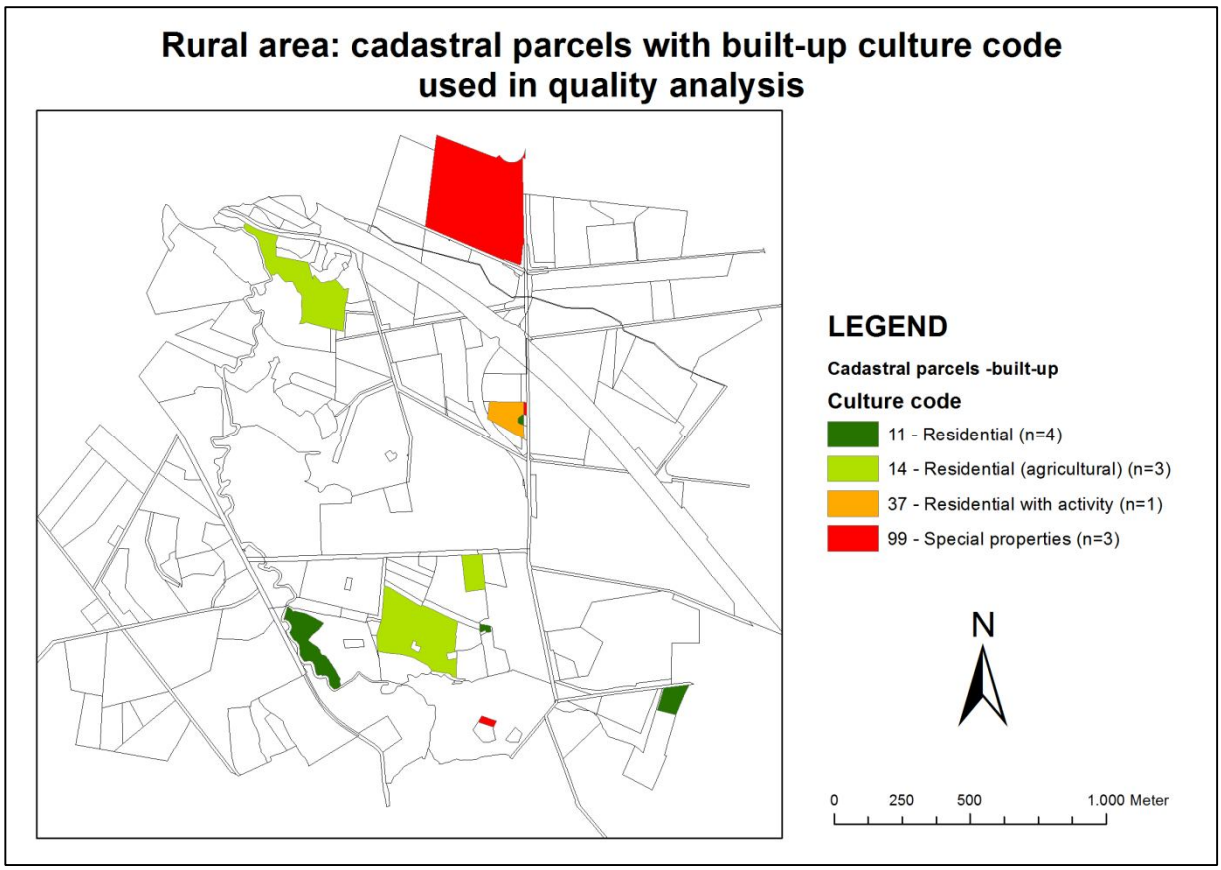


Figure 5.11: Cadastral parcels with a built-up culture code used in the quality analysis for the rural test area.

The results for the quality analysis on the built-up culture codes in the rural test area are presented in table 5.2 and in appendix D. In the table, the column "True" presents the number of selected cadastral parcels for which the calculated culture code is equal to the real-life situation. The column "False" gives the number of cadastral parcels for which the calculated culture code is not right. Finally, the total number of selected cadastral parcels per culture code is presented in the column "Total".

Culture code	True	False	Total
11 Residential	4	0	4
14 Residential (agricultural)	3	0	3
37 Residential with activity	1	0	1
99 Special properties	0	3	3
TOTAL	***8***	***3***	***11***

Table 5.2: Results of the quality analysis for built-up culture codes in the rural test area.

The table shows that the results for built-up culture codes in the rural test area are good, except for the built-up culture code "Special properties". The reason for the wrong results differs between the three selected cadastral parcels:"

- One cadastral parcels should have the culture code "Activity (utility)". However, as the main buildings of the utility complex are located on another, bordering, parcel, the selected parcel has not been given the right culture code. The cadastral parcel did not contain any indicators for the presence of a utility complex.
- One cadastral parcel should have the culture code "Residential". As the building is located on two cadastral parcels, the residential object with the purpose of use "residential function" has been joined with only one of these cadastral parcels. The selected cadastral parcel does not have a purpose of use located on it and, as a result, the cadastral parcel is given the culture code "Special properties". Figure 5.12 visualizes this problem.

Figure 5.12: Cadastral parcel without purpose of use (red) is given the built-up culture code "Special properties". The stars visualize the residential objects with purpose of use and the pink polygons visualize the BAG buildings.

- One cadastral parcel should have the culture code "Activity (agricultural)". This cadastral parcel contains BAG buildings without a residential object, as they are secondary buildings belonging to an agricultural main building. As no residential object (and, therefore, no purpose of use) is located on the parcel, the parcel is given the culture code "Special properties".

Results for non-built-up culture codes

Figure 5.13 shows which cadastral parcels with a non-built-up culture code have been selected for the quality analysis. Table 5.3 and appendix D present the results of this quality analysis for the different non-built-up culture codes present in the rural test area.

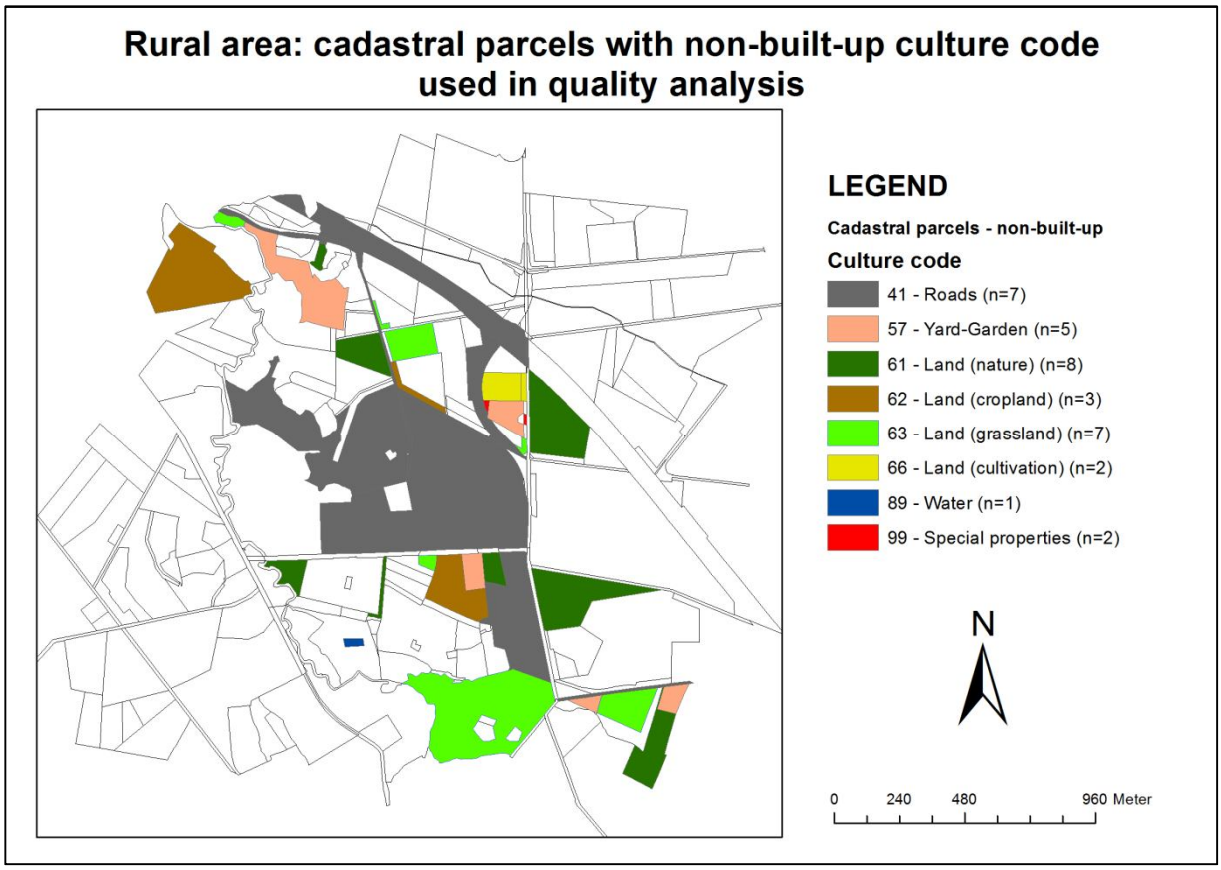


Figure 5.13: Cadastral parcels with a non-built-up culture code used in the quality analysis for the rural test area.

Culture code	True	False	Total
41 Roads	4	3	7
57 Yard-Garden	4	1	5
61 Land (nature)	8	0	8
62 Land (cropland)	2	1	3
63 Land (grassland)	7	0	7
66 Land (cultivation)	2	0	2
89 Water	---	---	---
99 Special properties	0	2	2
TOTAL	***27***	***7***	***34***

Table 5.3: Results of the quality analysis for non-built-up culture codes in the rural test area.

Again, the results of the method proposed in this document, are quite satisfying. However, for seven cadastral parcels a wrong non-built-up culture code has been calculated. Below, the wrong results are analyzed per culture code:

- Roads. The wrong results for this culture code can be explained by the fact that roads are given priority over other land use types such as cropland, grassland and nature. The percentage of the parcel covered by a road is not taken into account. In some cases, large cadastral parcels contain a road and have, therefore, been given the culture

code “Roads”. However, when comparing this result with the real-life situation, one might conclude that other land use types are more prominent at this parcel. An example is presented in figure 5.14. In this figure the large grey cadastral parcel has been given the culture code “Roads”, although only a small road line element is crossing the parcel. The culture code “Land (nature)” would have been a better result for the cadastral parcel.

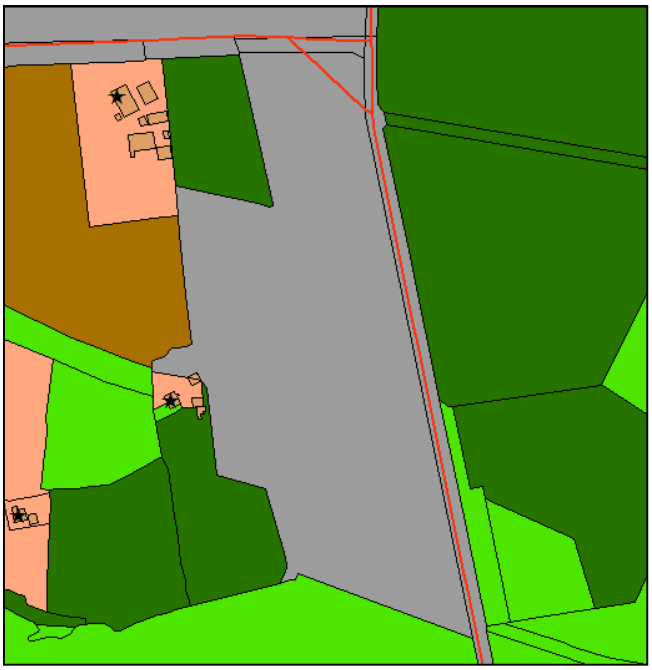


Figure 5.14: Road line element crossing a cadastral parcel which is dominated by the land use type “Land (nature).

- Yard-Garden. In the decision rules it is stated that cadastral parcels having a built-up culture code “Residential (agriculture)” and also requiring a non-built-up culture code, should get the non-built-up culture code “Yard-Garden”. This rule gives the right culture code for four parcels. However, for one, larger cadastral parcels this is not the right culture code. Rather, this cadastral parcel should have the non-built-up culture code “Land (grassland)”, as large parts of the parcel are used as grassland for cattle.
- Land (cropland). In this case the source dataset (TOP10NL) gives the wrong information. In the TOP10NL dataset the cadastral parcel has the terrain type “cropland”, whereas in the real-life situation only grassland is located on this parcel.
- Special properties. Both cadastral parcels do not have a built-up cadastral parcel and the TOP10NL terrain type located on the parcels is “Other”. As no other indicators are located on these cadastral parcels, they have been given the culture code “Special properties”. In the real-life situation one cadastral parcel is part of a garden belonging to a residential building, whereas the other contains forest.

In table 5.3 no results are presented for the culture code “Water”. Only one cadastral parcel in the rural test area has this non-built-up culture code. The author wanted to have a look at this parcel too, but it was not visible from public roads (because the view was blocked by a forest) and it was not possible to come closer to the parcel.

5.2 – POC: city centre

In this section, the case study for the second test area, the city centre, is presented. The method used is, to a large extent, the same as the one used in section 5.1. Therefore, only a detailed description of the procedure is given when this differs from the method in section 5.1.

5.2.1 Case study

Built-up, non-built-up or combined

Unlike in the rural area, the city centre also contains cadastral parcels of which more 90% is covered by BAG building polygons. This means that also the "*bebouwingscode*" 3 (only built-up) can be found in this test area. The different "*bebouwingscodes*" are given to the following numbers of cadastral parcels:

1. Only non-built-up: 78 parcels
2. Built-up and non-built-up: 200 parcels
3. Only built-up: 173 parcels

Figure 5.15 presents the map showing the distribution of these *bebouwingscodes* in the city centre test area.

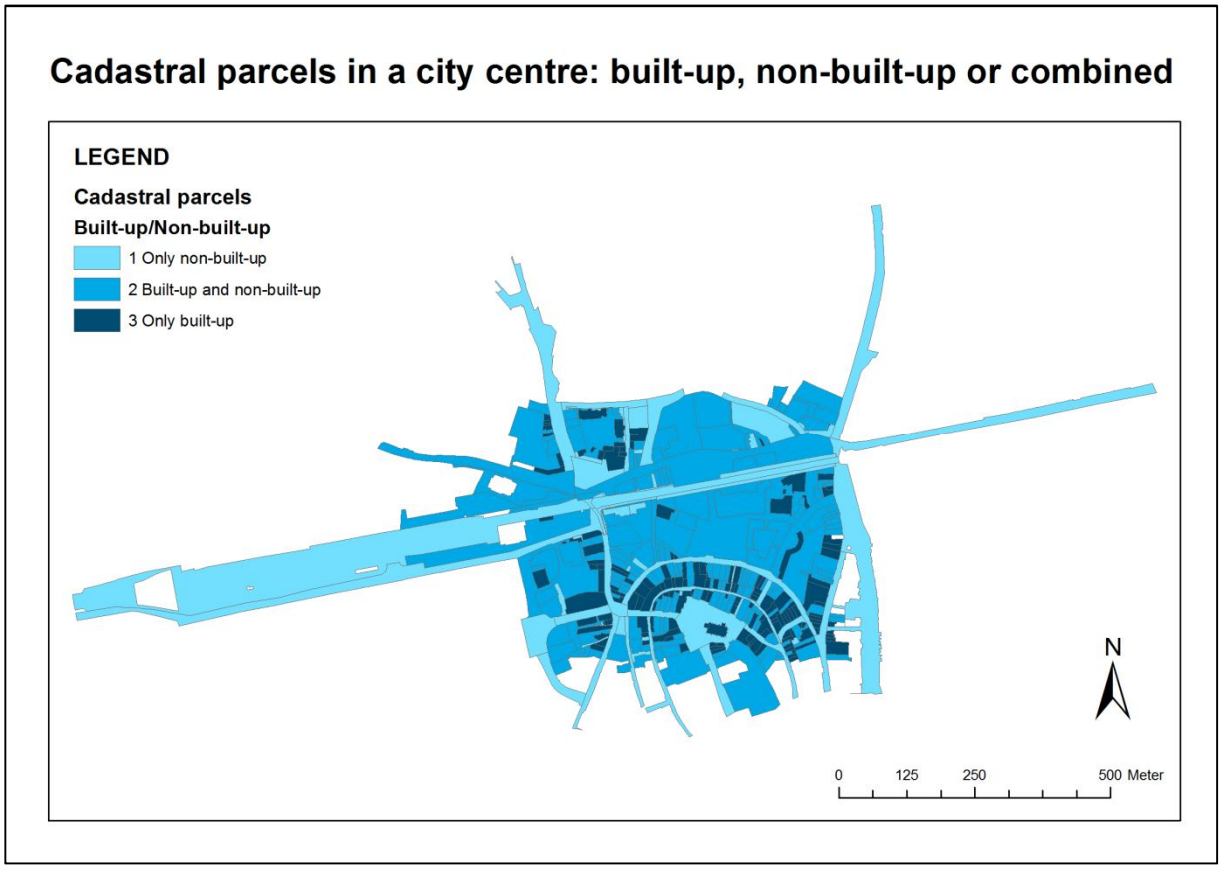


Figure 5.15: Cadastral parcels in a city centre area: built-up, non-built-up or combined.

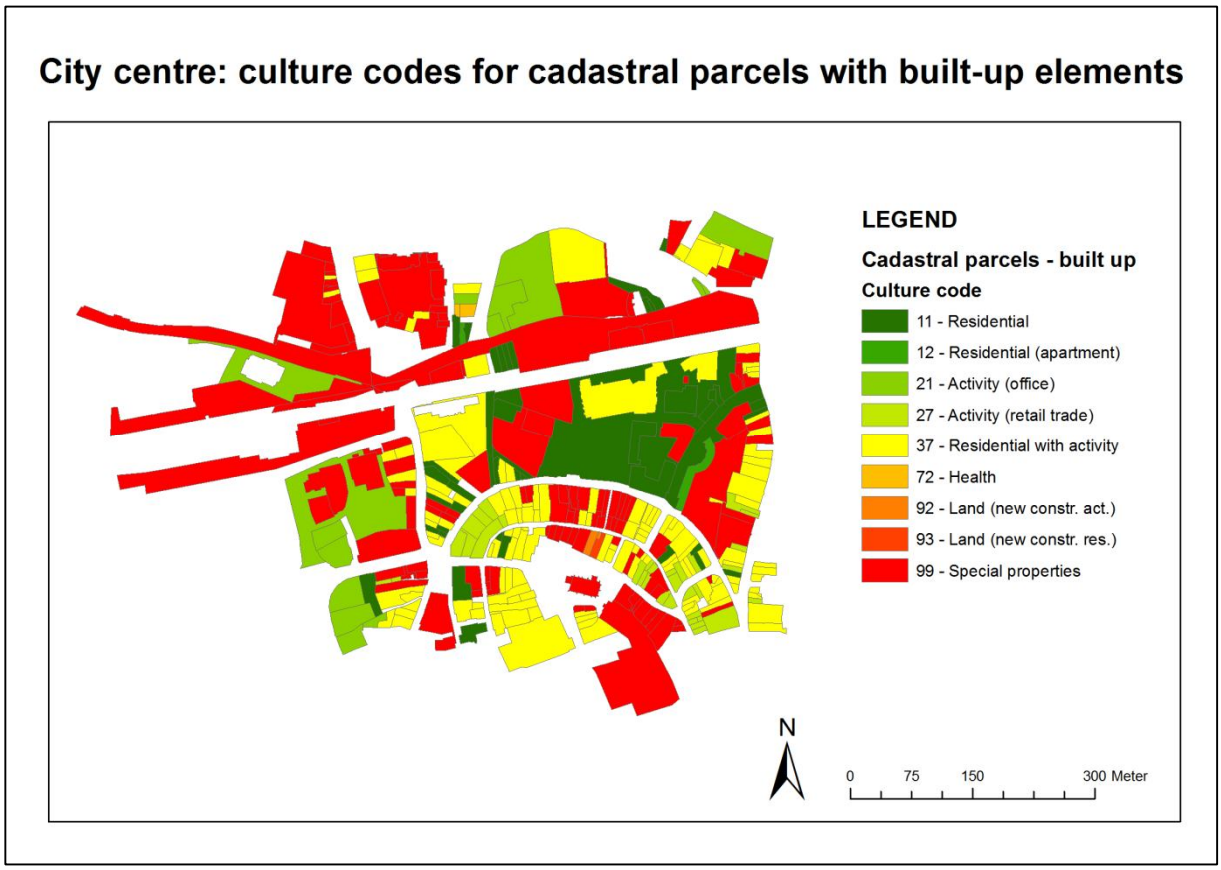


Figure 5.16: Built-up culture codes in the city centre.

Built-up culture codes

As the number of cadastral parcels requiring a built-up culture code and the number of residential objects are much larger than in the rural area, it would be a very time-consuming activity to use the information button in ArcGIS for every residential object, as has been done for the rural test area. Therefore, another method was used in the city centre test area. An individual layer was created for every purpose of use from the BAG dataset. All these layers were joined to the cadastral parcel layer by making use of the Spatial join tool in ArcGIS. This resulted in an attribute table containing a number of different columns, all giving information on the presence of specific purposes of use. With this information it was, for a number of cadastral parcels, possible to define which built-up culture should be given to the cadastral parcel. For the other parcels the TOP10NL has to be used or they are given the built-up culture code “Special properties”. Unlike in the rural test area, the city centre test area does include cadastral parcels with a residential function and apartment rights located on them. Therefore, two cadastral parcels are given the built-up culture code “Residential (apartment)” instead of “Residential”. Furthermore, in the city centre three BAG buildings had either the status “building permit granted” or “building process started”. As a result, two cadastral parcels are given the built-up culture code “Land (new construction activity)” and one was given the built-up culture code “Land (new construction residential)”. Figure 5.16 presents a map showing the distribution of the built-up culture codes in the city centre test area.

Non-built-up culture codes

In the city centre test area, like in the rural area, no built-up culture codes “Activity (industry)” and “Defense” were present. However, 106 cadastral parcels are given the non-built-up culture code “Yard – Garden”, as these parcels possessed either the built-up culture code “Residential” or the built-up culture code “Residential with activity”.

Using the object classes from the TOP10NL gives the following results:

- Functional area. Not present in the city centre area.
- Railroad section. 6 Cadastral parcels are given the non-built-up culture code "Public transport".
- Road section. First, the road section polygons were used for giving the culture code "Parking" to 11 cadastral parcels. Second, the road section central line was used, which resulted in 36 cadastral parcels being given the non-built-up culture code "Roads".
- Cultural landscape objects. Not present in the city centre area.
- Terrain classes and water objects. This gives the following results:
    - Water: not present in city centre area.
    - Land (nature): never the highest percentage.
    - Land (grassland): 1 cadastral parcel.
    - Funeral service: never the highest percentage.

The remaining 118 cadastral parcels still requiring a non-built-up culture code are given the non-built-up culture code "Special properties". Figure 5.17 presents a map showing the distribution of the different non-built-up culture codes in the city centre test area and table 5.4 gives the numbers of cadastral parcels assigned to the different built-up and non-built-up culture codes.

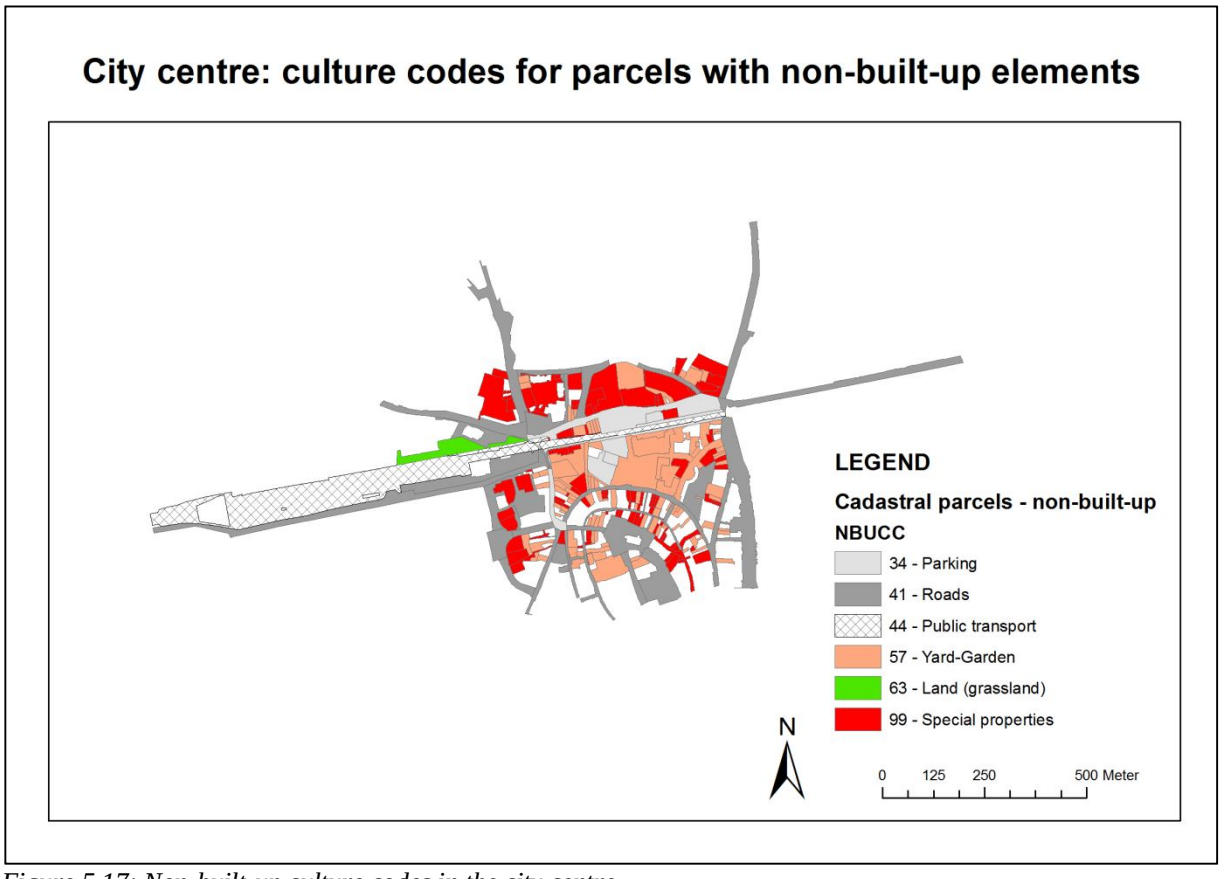


Figure 5.17: Non-built-up culture codes in the city centre.

Culture code	Built-up	Non-built-up
11 Residential	45	
12 Residential (apartment)	2	
21 Activity (office)	15	
27 Activity (retail trade)	33	
34 Parking		11
37 Residential with activity	135	
41 Roads		36
44 Public transport		6
57 Yard – Garden		106
63 Land (grassland)		1
72 Health	2	
92 Land (new construction activity)	2	
93 Land (new construction residential)	1	
99 Special properties	138	118
TOTAL	***451***	***278***

Table 5.4: Results for built-up and non-built-up culture codes in city centre.

5.2.2 Quality analysis

Just like for the rural test area, the results for the city centre test area are validated by means of a quality analysis. The same quality analysis method as in section 5.1.2 has been applied in this section.

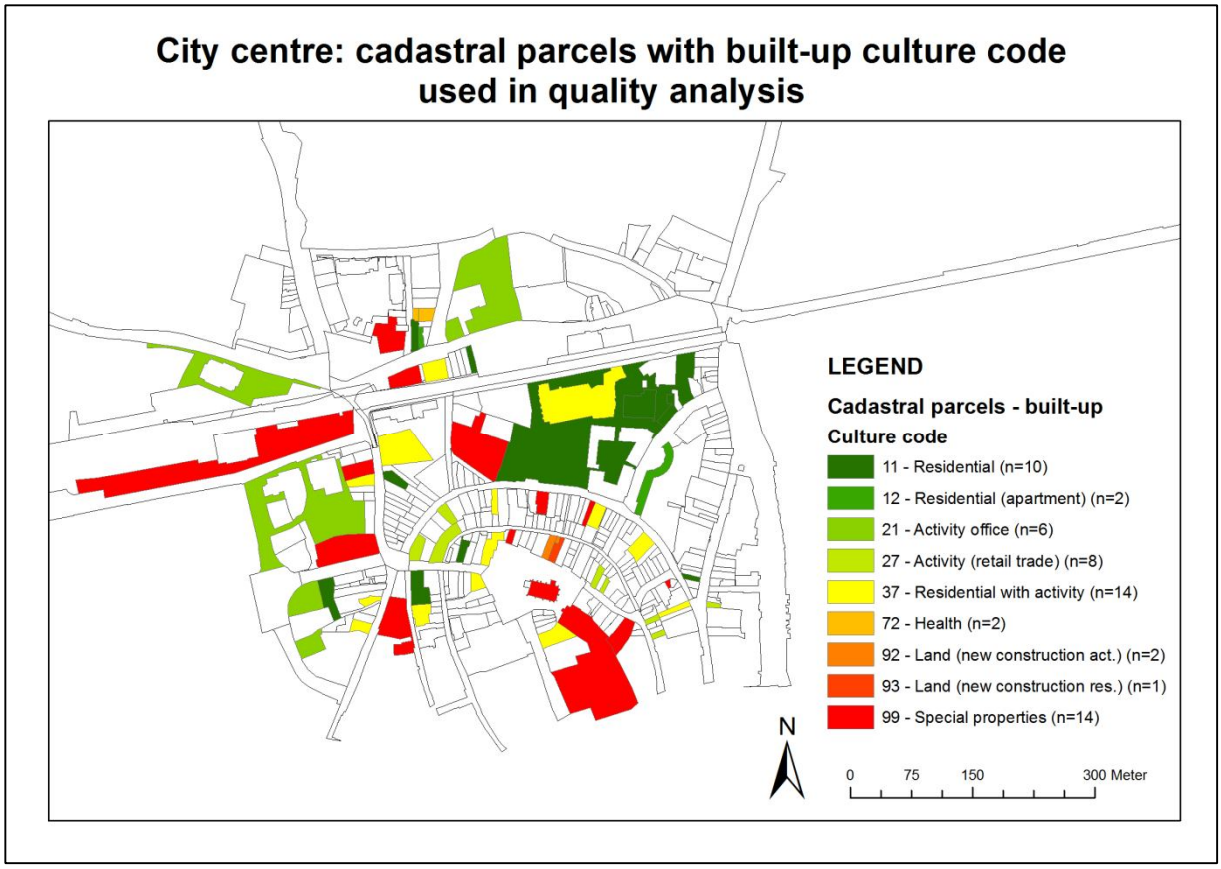


Figure 5.18: Cadastral parcels with a built-up culture code used in the quality analysis for the city centre test area.

Results for built-up culture codes
Figure 5.18 presents a map showing the cadastral parcels with a built-up code which are selected for the quality analysis.

Culture code	True	False	Total
11 Residential	4	6	10
14 Residential (apartment)	2	0	2
21 Activity (office)	6	0	6
27 Activity (retail trade)	8	0	8
37 Residential with activity	12	2	14
72 Health	2	0	2
92 Land (new construction act.)	*2*	*0*	*2*
93 Land (new construction res.)	*0*	*1*	*1*
99 Special properties	6	8	14
TOTAL	***40***	***16***	***56***

Table 5.5: Results of the quality analysis for built-up culture codes in the city centre test area.

Table 5.5 and appendix E present the results of the quality analysis for built-up culture codes in the city centre test area. The results for the culture codes "Land (new construction activity)" and "Land (new construction residential)" have not been taken into account in the total numbers of wrong and right results, as the construction works were not visible for the author. However, the author is not sure about the culture code "Land (new construction residential)", as the current building is used for catering purposes and does not have a residential function.

For most culture codes the results are satisfying, except those of the culture codes "Residential" and "Special properties". The reasons for the large number of wrong results for these culture codes are the following:

- Residential. Five out of six wrong results can, unfortunately, be explained by a mistake in the calculations. As the culture codes have been manually assigned to the cadastral parcels, a number of mistakes has been made concerning the built-up culture code "Residential" in the city centre test area. Re-analyzing the source dataset's contents, shows that three of these cadastral parcels should have the culture code "Activity (retail trade)", that one should get the culture code "Special properties", and one the culture code "Health". The wrong result of the remaining cadastral parcel is explained by a wrong purpose of use in the BAG dataset. Instead of "residential function" and "retail trade function", only the residential function is present in the BAG dataset.
- Special properties. The main reason why there are so many wrong results for this culture code is caused by the lacking (spatial) information on the purposes of use "catering" and "religion". The BAG dataset does not contain information on catering and religion. Rather, restaurants, bars, churches, etc. are given the purpose of use "gathering function". In the method proposed in this document, these purposes of use cause, when no additional information is available in the TOP10NL dataset, the cadastral parcels to have the culture code "Special properties". Six out of eight wrong results can be explained by this reason; five parcels should have the culture code "Activity (catering)", one the culture code "Religion". One mistake is caused by a wrong purpose of use in the BAG dataset. This parcel should have the purpose of use "office function" and the culture code "Activity (office)". The last wrong result is caused by the mistake explained in the previous bullet. The author has assigned the

cadastral parcel the culture code “Special properties”, although the parcel contains the purpose of use “residential function”.

Without the mistake made by the author during the calculations, the calculations would have given only ten wrong results. This means the method proposed, when applied without mistakes, would give even more satisfying results than it has given now.

Results for non-built-up culture codes
Figure 5.19 shows the cadastral parcels with a non-built-up culture code selected for the quality analysis, and table 5.6 and appendix E present the results of the quality analysis.

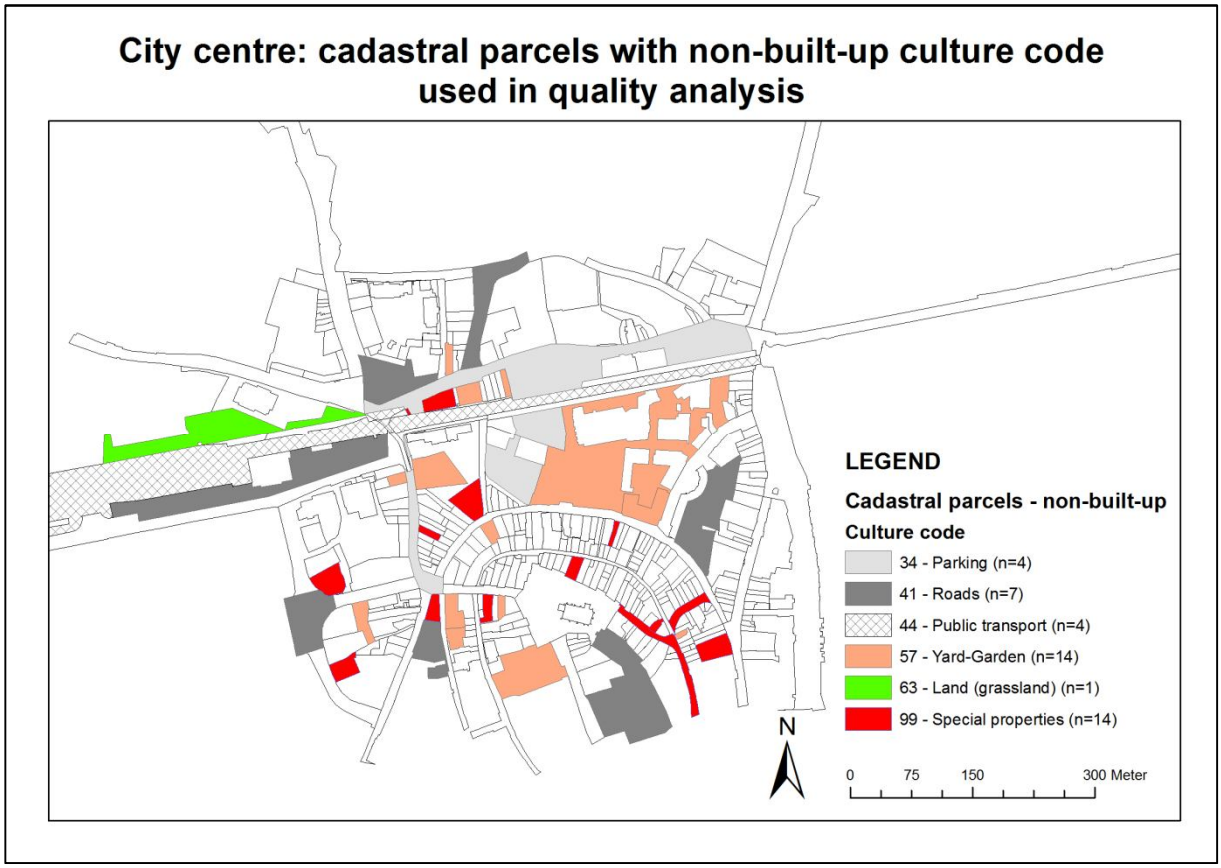


Figure 5.19: Cadastral parcels with a non-built-up culture code used in the quality analysis for the city centre test area.

Culture code	True	False	Total
34 Parking	2	2	4
41 Roads	6	1	7
44 Public transport	4	0	4
57 Yard-Garden	10	4	14
63 Land (grassland)	1	0	1
99 Special properties	13	1	14
TOTAL	***36***	***8***	***44***

Table 5.6: Results of the quality analysis for non-built-up culture codes in the city centre test area.

Most of the wrong results are found for the non-built-up culture codes “Parking” and “Yard-Garden”. In the case of the culture code “Yard-Garden”, the wrong results can be explained by the mistake made by the author. Some cadastral parcels are given a residential function,

although this function is not present at the parcel. As a result, these cadastral parcels are given the non-built-up culture code “Yard-Garden”. When re-analyzing the information from the source datasets, these cadastral parcels should have been given the non-built-up culture code “Special properties”.

The wrong results for the culture code “Parking” are caused by the priority given to parking. Parking has a higher priority than, for example, roads. This means that when parking is located on a cadastral parcel, this parcel will get the culture code “Parking”, although in some cases roads are dominating the appearance of this parcel.

Again, like for the results of the built-up culture codes, without the mistake made during the calculations, the results of the case study would have been even more satisfying. When re-calculating, only four wrong results are found for the selected cadastral parcels with a non-built-up culture code.

5.3 – POC: living area

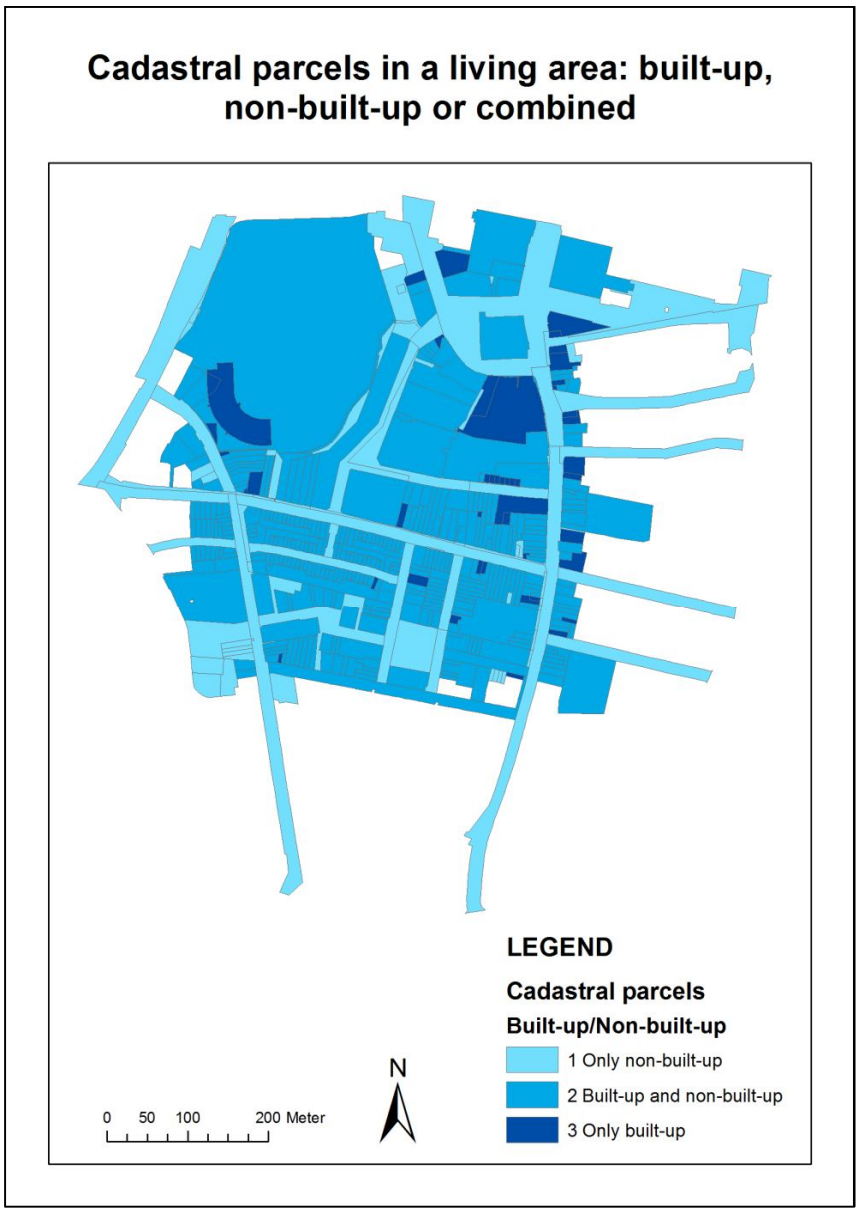


Figure 5.20: Cadastral parcels in a living area: built-up, non-built-up or combined.

The same testing procedure as in section 5.2 is applied to a living area test site. Only when the testing procedure differs from the one applied in section 5.2, more information is given on the testing method.

5.3.1 Case study

Built-up, non-built-up or combined

In the living area, defining whether a cadastral parcel should get a built-up culture code, a non-built-up culture code or both gives the following outcome (also visualized by the map in figure 5.20 on the previous page):

1. Only non-built-up: 93 cadastral parcels.
2. Built-up and non-built-up: 386 cadastral parcels.
3. Only built-up: 47 cadastral parcels.

Built-up culture codes

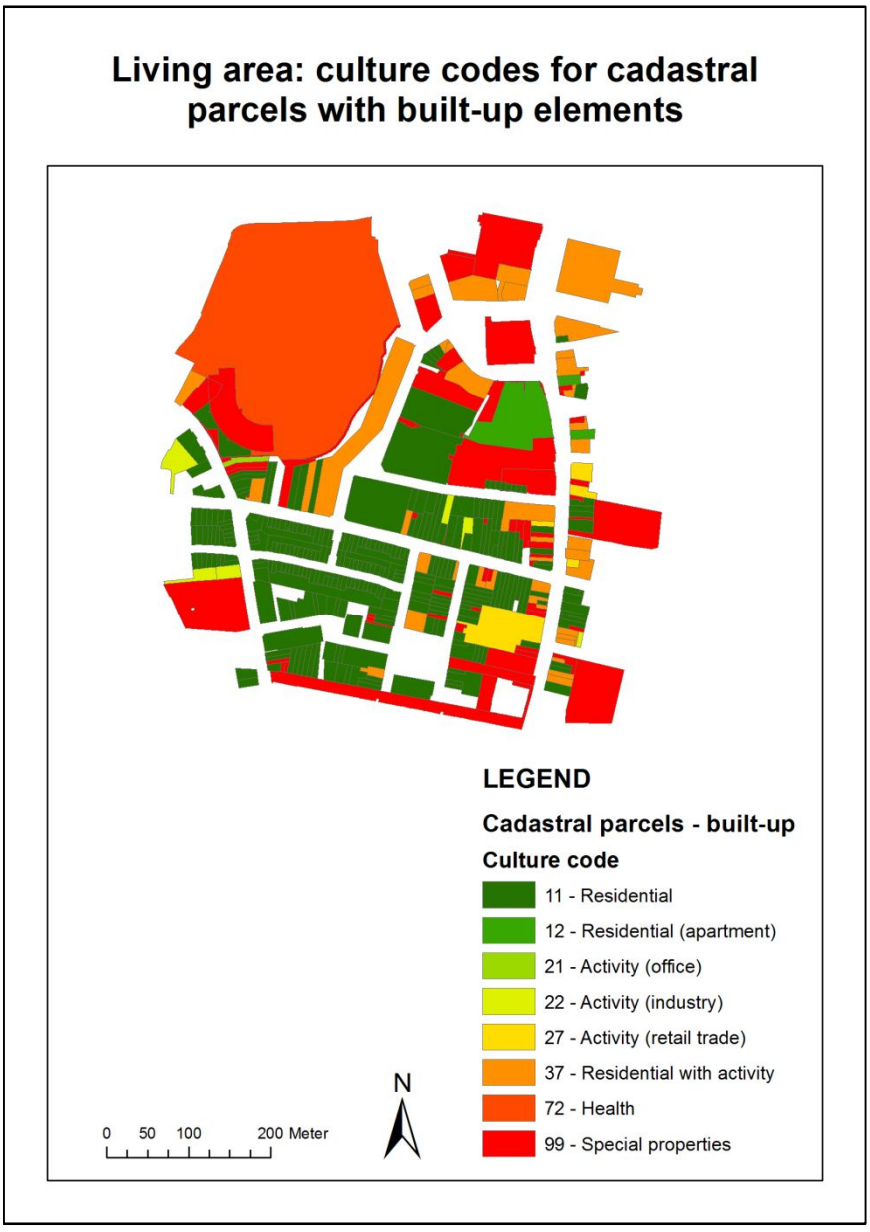


Figure 5.21: Built-up culture codes in a living area.

In this test area, four cadastral parcels are given the built-up culture code “Residential (apartment)”. In the living area no BAG buildings have the status “building process started” or “building permit granted”. Therefore, no cadastral parcels are given the built-up culture codes “Land (new construction activity)” or “Land (new construction residential)”. Furthermore, no built-up culture codes are given to cadastral parcel based on the TOP10NL object classes. Finally, 69 cadastral parcels are given the built-up culture code “Special properties”. In figure 5.21, the map shows the distribution of the different built-up culture codes in the living area.

Non-built-up culture codes

Unlike in the previous two test areas, in the living area the built-up culture code “Activity (industry)” can be found. As a result, 6 cadastral parcels are given the non-built-up culture code “Land (industry)”. Furthermore, 320 cadastral parcels are given the non-built-up culture code “Yard – Garden”, as these parcels possess the built-up culture codes “Residential” or “Residential with activity”.

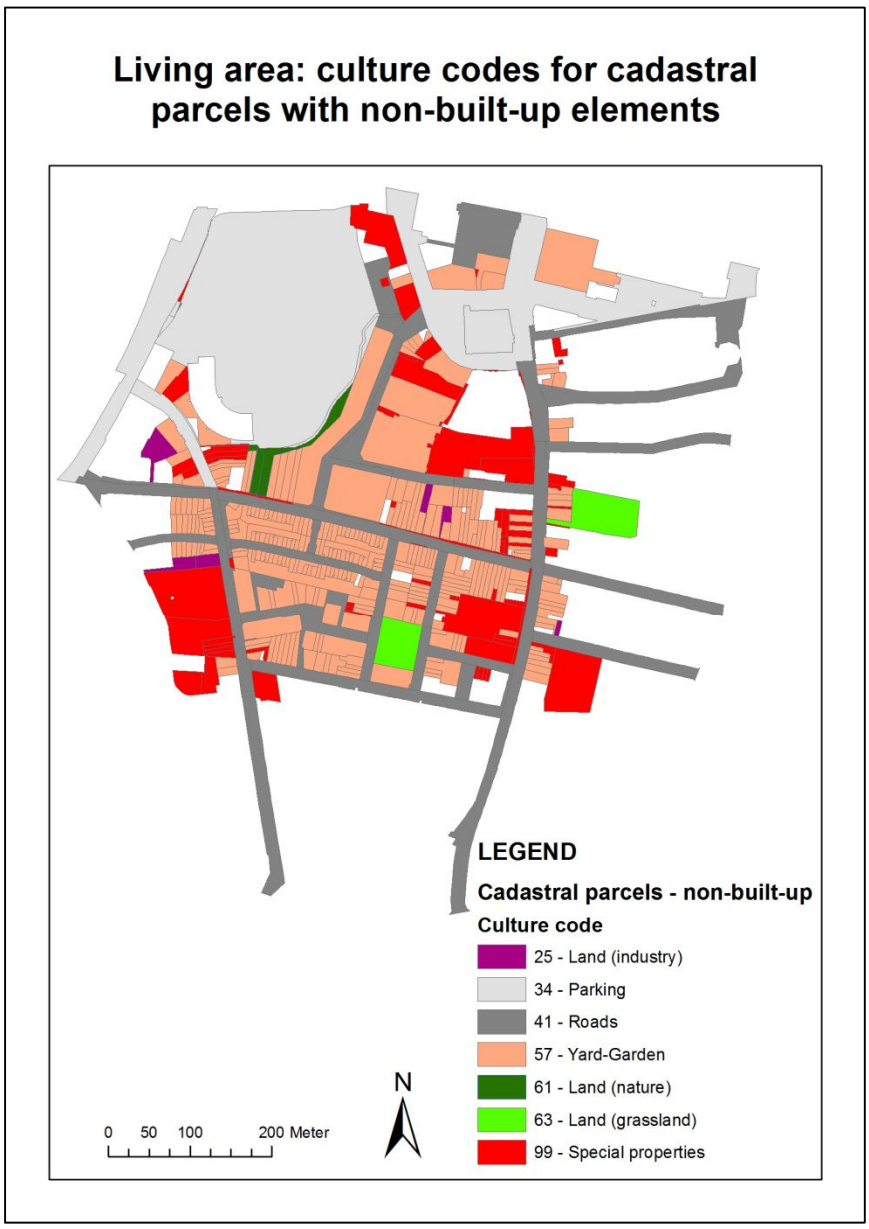


Figure 5.22: Non-built-up culture codes in a living area.

Making use of the TOP10NL object classes gives the following results for the non-built-up culture codes in the living area:

- Functional area. Not present in the living area.
- Railroad section. Not present in the living area.
- Road section. First, by using the road section polygons 10 cadastral parcels are given the non-built-up culture code "Parking". Second, by using the road section central lines 31 cadastral parcels are given the non-built-up culture code "Roads".
- Cultural landscape objects. Not present in the living area.
- Terrain classes and water objects. This gives the following results:
    - Water: 0 cadastral parcels (not present)
    - Land (nature): 4 cadastral parcels
    - Land (grassland): 1 cadastral parcel

The remaining 105 cadastral parcels are given the non-built-up culture code "Special properties". In figure 5.22, the distribution of the non-built-up culture codes over the living area is presented. Finally, table 5.7 gives the numbers of cadastral parcels assigned to the different built-up and non-built-up culture codes.

Culture code	Built-up	Non-built-up
11 Residential	299	
12 Residential (apartment)	4	
21 Activity (office)	1	
22 Activity (industry)	6	
25 Land (industry)		6
27 Activity (retail trade)	5	
34 Parking		10
37 Residential with activity	41	
41 Roads		31
57 Yard – Garden		320
61 Land (nature)		4
63 Land (grassland)		2
72 Health	1	
99 Special properties	76	105
TOTAL	***433***	***478***

Table 5.7: Results for built-up and non-built-up culture codes in the living area.

5.3.2 Quality analysis

Results for built-up culture codes

Figure 5.23 presents a map showing the cadastral parcels with a built-up culture code which have been selected for the quality analysis. Table 5.8 and appendix F present the results of the quality analysis.

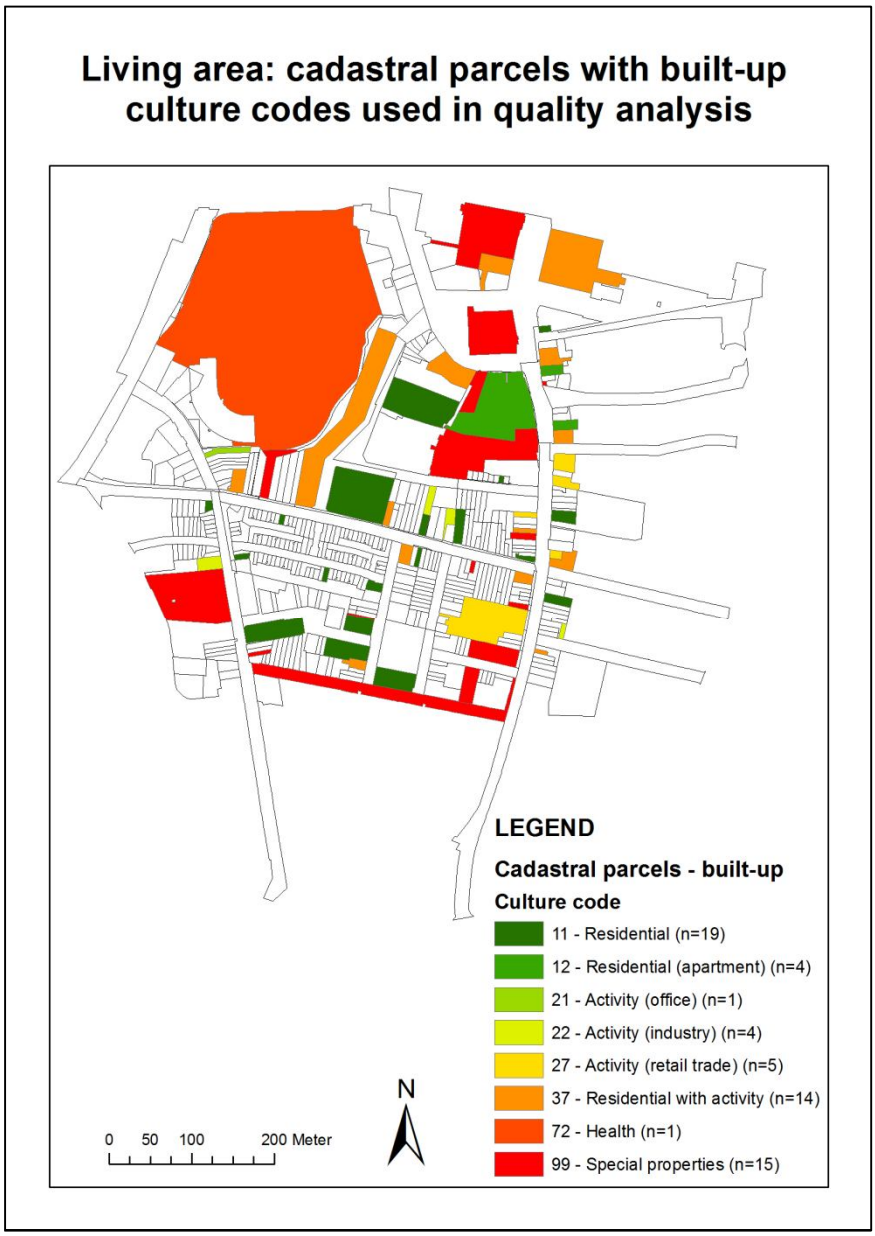


Figure 5.23: Cadastral parcels with a built-up culture code used in the quality analysis for the living area test area.

Culture code	True	False	Total
11 Residential	15	4	19
12 Residential (apartment)	4	0	4
21 Activity (office)	1	0	1
22 Activity (industry)	4	0	4
27 Activity (retail trade)	5	0	5
37 Residential with activity	14	0	14
72 Health	1	0	1
99 Special properties	4	11	15
TOTAL	***48***	***15***	***63***

Table 5.8: Results of the quality analysis for built-up culture codes in the living area test area.

In the living area, the only wrong results for the built-up culture codes are calculated for the culture codes “Residential” and “Special properties”. The reasons for these wrong results are the following:

- Residential. Two wrong culture codes are caused by the fact that the dataset with apartments is not complete. At two cadastral parcels apartments are located, but as this information is not available in the source dataset, the purpose of use “residential function” from the BAG dataset is used to give these parcels the built-up culture code “Residential”. The other two wrong culture codes are caused by missing purposes of use in the BAG dataset. At these cadastral parcels a combination of residential and retail trade units is present. However, the purpose of use “retail trade function” from the BAG dataset is missing for these parcels.
- Special properties. Nine out of eleven wrong results can be explained by the fact that BAG buildings are divided by parcel boundaries, and – as has been explained in section 5.1 on the rural area – that the residential objects are not located on these nine parcels. This means that the cadastral parcels do not contain a residential object and purpose of use. As no additional information is delivered by the TOP10NL dataset, these parcels are given the culture code “Special properties”. The final two wrong results are caused by the fact that BAG buildings are crossing parcel boundaries and, as a result, cause parcels dominated by roads to have a built-up culture code too. The buildings might have a wrong geometry, it might be caused by differences in production method between datasets, or it might be the real situation. In the latter case, the culture code is not wrong.

Results for non-built-up culture codes

Figure 5.24 shows the cadastral parcels with a non-built-up culture code used in the quality analysis, and table 5.9 and appendix F present the results of the quality analysis.

Culture code	True	False	Total
25 Land (industry)	4	0	4
34 Parking	4	2	6
41 Roads	7	0	7
57 Yard-Garden	21	1	22
61 Land (nature)	2	0	2
63 Land (grassland)	2	0	2
99 Special properties	10	2	12
TOTAL	***50***	***5***	***55***

Table 5.9: Results of the quality analysis for non-built-up culture codes in the living area test area.

The results for the non-built-up culture codes in the living area are very satisfying. Only in the case of the culture code “Parking” a relatively high number of wrong results is found. These wrong results can be explained by the priority given to parking, as has been explained in section 5.2.2.

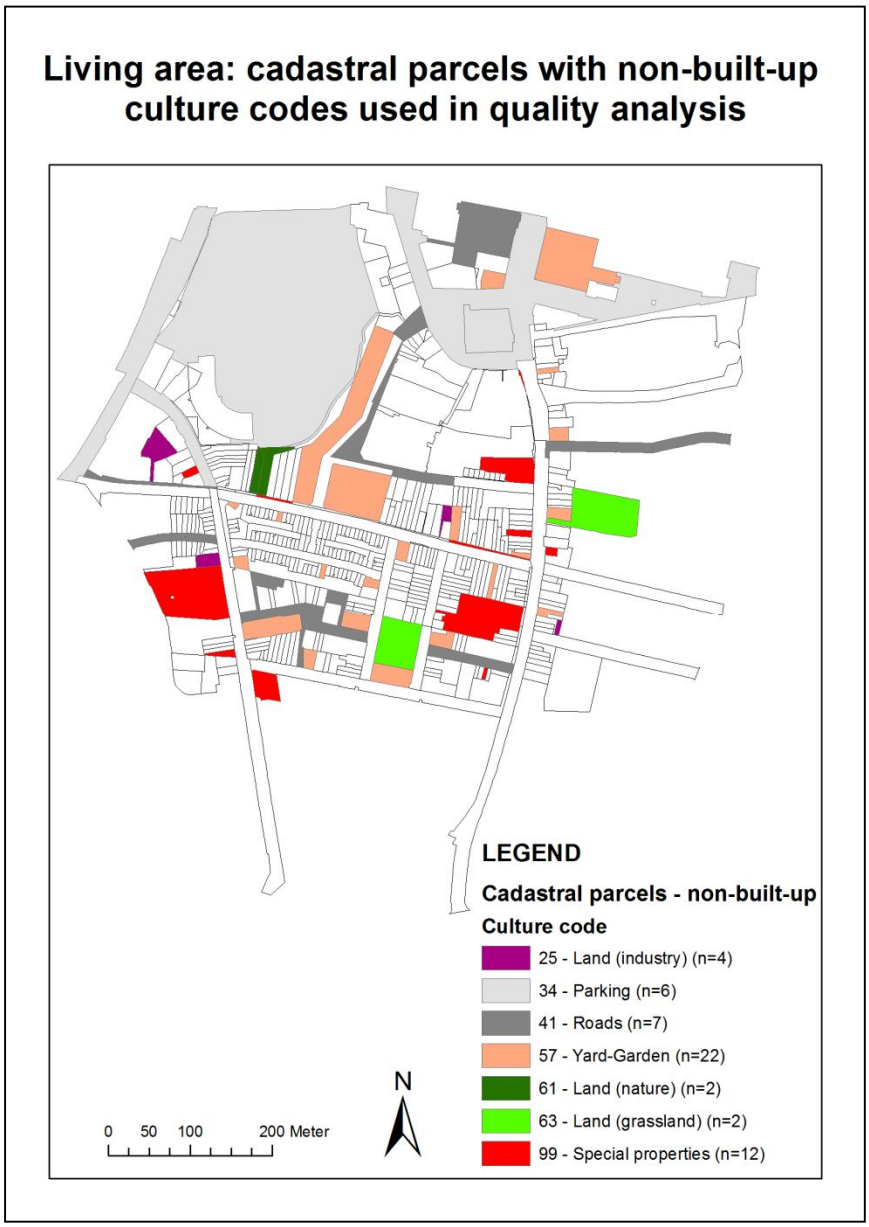


Figure 5.24: Cadastral parcels with a non-built-up culture code used in the quality analysis for the living area test area.

5.4 – POC: industrial/harbor area

The same testing procedure as in section 5.2 is applied to an industrial/harbor area test site. Only when the testing procedure differs from the one applied in section 5.2, more information is given on the testing method.

5.4.1 Case study

Built-up, non-built-up or combined

In the industrial/harbor area, defining whether a cadastral parcel should get a built-up culture code, a non-built-up culture code or both gives the following outcome (also visualized by the map in figure 5.25):

1. Only non-built-up: 60 cadastral parcels.
2. Built-up and non-built-up: 127 cadastral parcels.
3. Only built-up: 10 cadastral parcels.

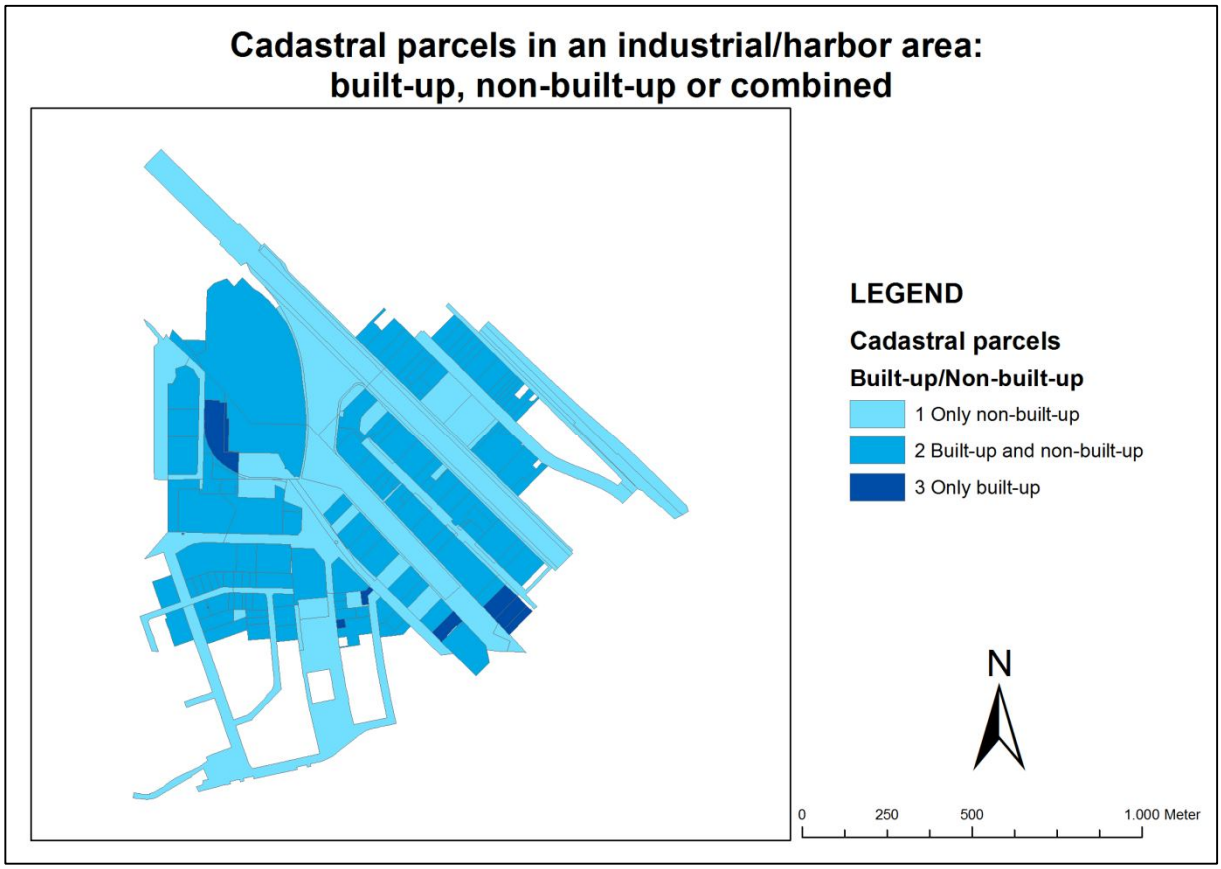


Figure 5.25: Cadastral parcels in an industrial/harbor area: built-up, non-built-up or combined.

Built-up culture codes

No apartment rights are located in the industrial/harbor test area. Therefore, no cadastral parcels are given the built-up culture code “Residential (apartment)”. Two cadastral parcels contain a BAG building with the status “building permit granted”. One of these parcels is given the built-up culture code “Land (new construction activity)” and one “Land (new construction residential)”. Furthermore, no built-up culture codes are given to cadastral parcel based on the TOP10NL object classes. Finally, 60 cadastral parcels are given the built-up culture code “Special properties”. In figure 5.26, the map shows the distribution of the different built-up culture codes in the industrial/harbor area.

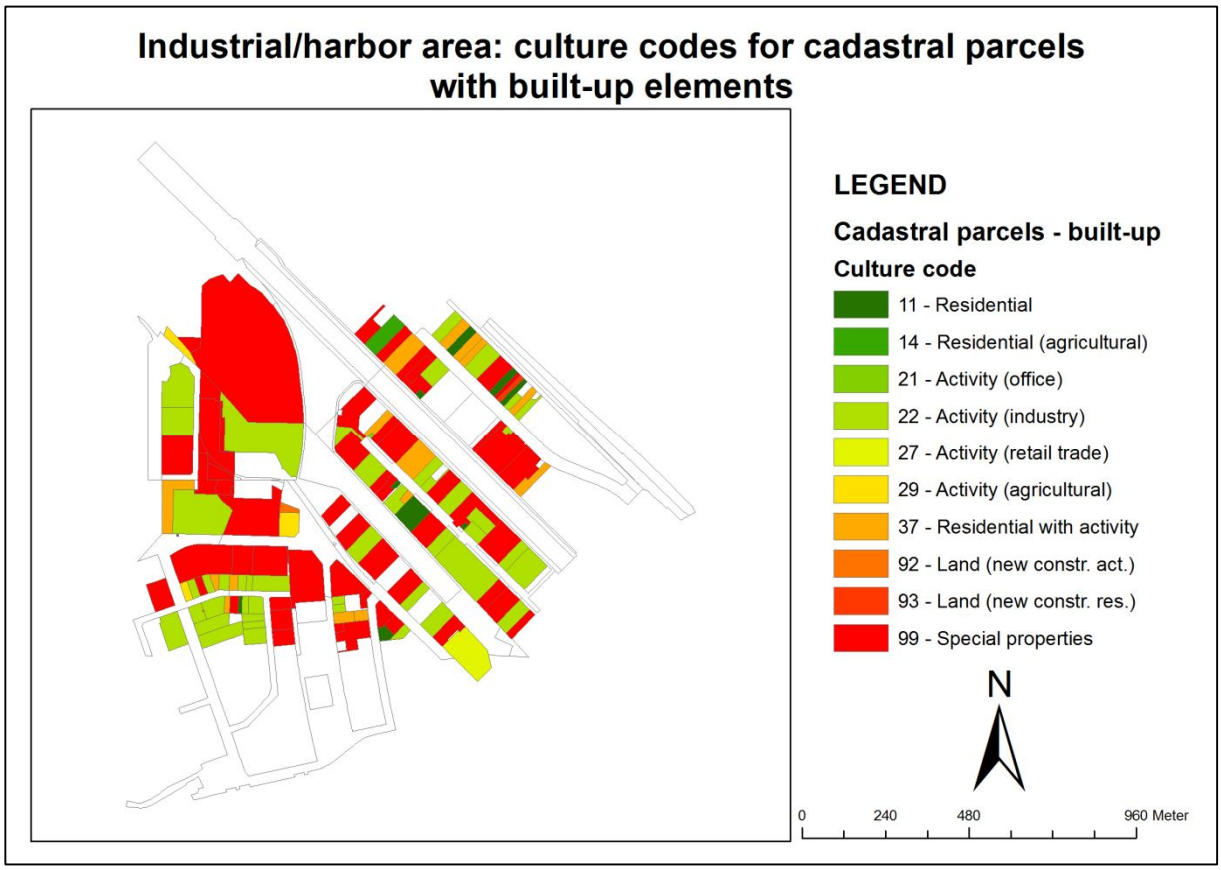


Figure 5.26: Built-up culture codes in an industrial/harbor area.

Non-built-up culture codes

In the industrial/harbor area, 44 cadastral parcels have the built-up culture code "Activity (industry)". These parcels are given the non-built-up culture code "Land (industry)". Furthermore, 25 cadastral parcels are given the non-built-up culture code "Yard – Garden", as these parcels possess the built-up culture codes "Residential", "Residential (agricultural)" or "Residential with activity".

Making use of the TOP10NL object classes gives the following results for the non-built-up culture codes in the industrial/harbor area:

- Functional area. Two cadastral parcels contain the functional area type "port". These two parcels are given the non-built-up culture code "Port".
- Railroad section. One cadastral parcels contains a railroad section and is given the non-built-up culture code "Public transport".
- Road section. The polygon road sections do not contain parking areas in the industrial/harbor area. As a result, no cadastral parcels are given the non-built-up culture code "Parking". By using the road section central lines 13 cadastral parcels are given the non-built-up culture code "Roads".
- Cultural landscape objects. In the industrial/harbor area, two cadastral parcels are intersected by a "cultural landscape objects" line element. These two cadastral parcels are given the non-built-up culture code "Port".
- Terrain classes and water objects. This gives the following results:"
    - Water: in this area, a distinction can be made between water objects with the function "port" and sections with the function "other".
        - Port: 7 cadastral parcels
        - Water: 1 cadastral parcel

- Land (nature): 14 cadastral parcels
- Land (grassland): 6 cadastral parcel

After using the terrain classes and water objects from the TOP10NL, 72 cadastral parcels which require a non-built-up culture code remain without one. These cadastral parcels are given the non-built-up culture code "Special properties". Figure 5.27 presents a map showing the distribution of the non-built-up culture codes over the industrial/harbor area. Furthermore, table 5.10 sums up the results for the industrial/harbor test area, by presenting the complete list of culture codes in the industrial/harbor area and the number of cadastral parcels which are given these culture codes.

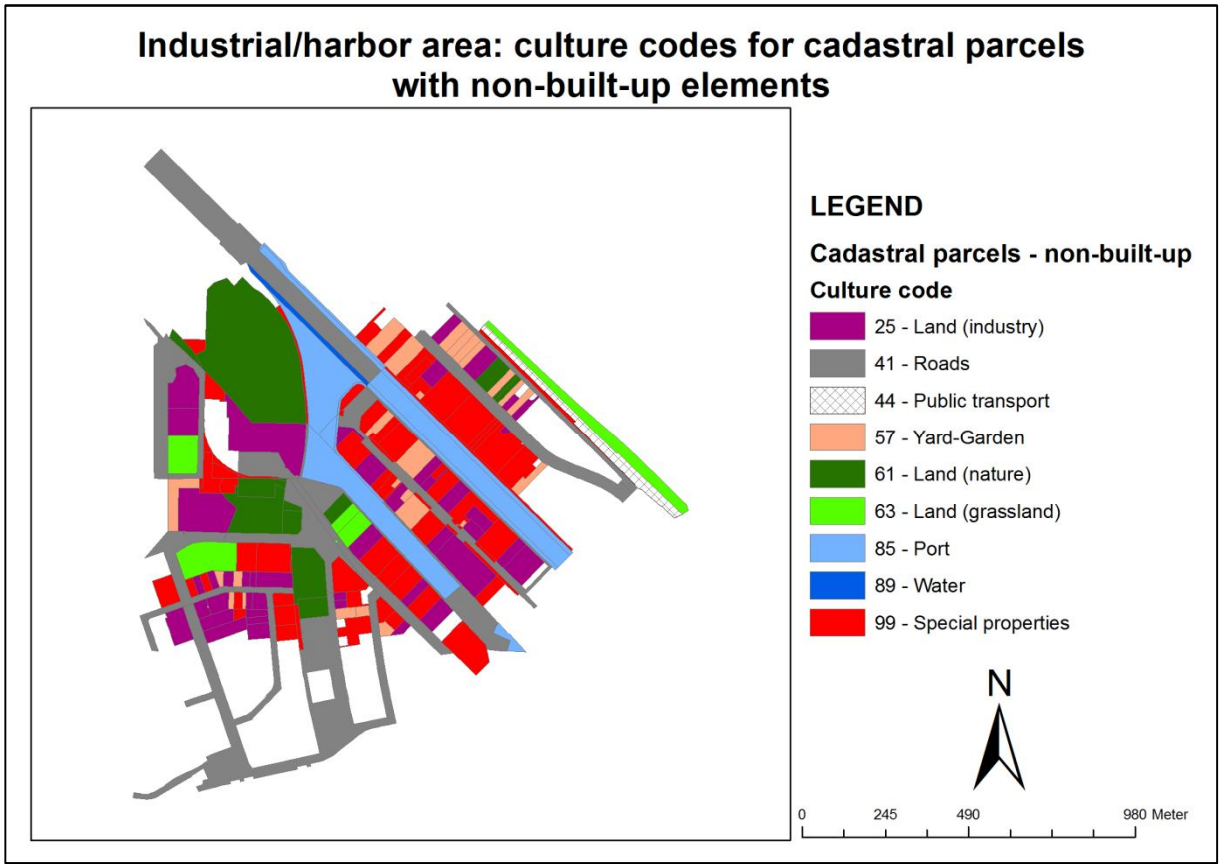


Figure 5.27: Non-built-up culture codes in an industrial/harbor area.

Culture code	Built-up	Non-built-up
11 Residential	9	
14 Residential (agricultural)	1	
21 Activity (office)	1	
22 Activity (industry)	45	
25 Land (industry)		44
27 Activity (retail trade)	1	
29 Activity (agricultural)	3	
37 Residential with activity	15	
41 Roads		13
44 Public transport		1
57 Yard – Garden		25
61 Land (nature)		14
63 Land (grassland)		6

85 Port		11
89 Water		1
92 Land (new construction activity)	1	
93 Land (new construction residential)	1	
99 Special properties	60	72
TOTAL	***137***	***187***

Table 5.10: Results for built-up and non-built-up culture codes in the industrial/harbor area.

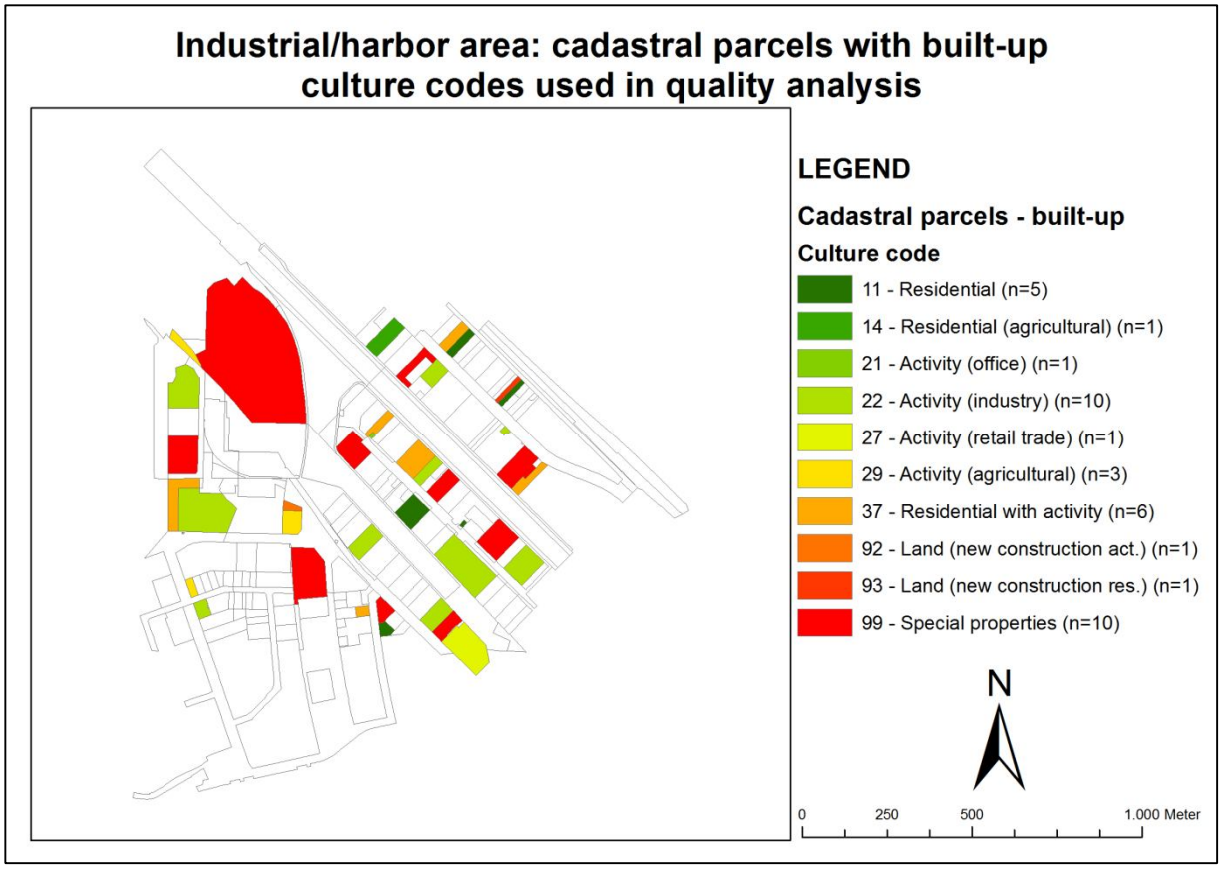


Figure 5.28: Cadastral parcels with a built-up culture code used in the quality analysis for the industrial/harbor test area.

Culture code	True	False	Total
11 Residential	4	1	5
14 Residential (agricultural)	0	1	1
21 Activity (office)	1	0	1
22 Activity (industry)	10	0	10
27 Activity (retail trade)	1	0	1
29 Activity (agricultural)	0	3	3
37 Residential with activity	5	1	6
92 Land (new construction act.)	*1*	*0*	*1*
93 Land (new construction res.)	*0*	*1*	*1*
99 Special properties	7	3	10
TOTAL	***28***	***9***	***37***

Table 5.11: Results of the quality analysis for built-up culture codes in the industrial/harbor test area.

5.4.2 Quality analysis

Results for built-up culture codes

Figure 5.28 presents a map showing the cadastral parcels with a built-up culture code which have been selected for the quality analysis. The results of the quality analysis are presented in table 5.11 and appendix G. Wrong results are found for the culture codes “Residential”, “Residential (agricultural)”, “Activity (agricultural)”, “Residential with activity” and “Special properties”. As the construction activities, on which the culture codes numbers 92 and 93 are based, were not visible from the outside, the author cannot really judge whether the culture code is right or wrong. Therefore, the results are shown in italics. However, the author has observed that in case of the culture code “Land (new construction residential)” there is currently no residential function present on the cadastral parcel. As this might change because of the construction activity one cannot state that the culture code is wrong. These culture codes are, for these reasons, not included in the totals in the last row of the table.

The wrong results for the built-up culture codes in the industrial/harbor test area can be explained by the following reasons:

- Residential. The wrong result is caused by incorrect information in the BAG dataset. At the cadastral parcel no residential function is observed by the author. Rather, the cadastral parcel contains a combination of retail trade and offices. This means the right built-up culture code for the cadastral parcel would be “Special properties”.
- Residential (agricultural). In this case, a cadastral parcel contains the purposes of use “residential function” and “industrial function”. As grassland is located within a distance of 50m from the parcel (see figure 5.29), the parcel is given the culture code “Residential (agricultural)”. However, the author has observed the right culture code for this parcel would have been “Residential with activity”, because no agricultural activity is found at the parcel.

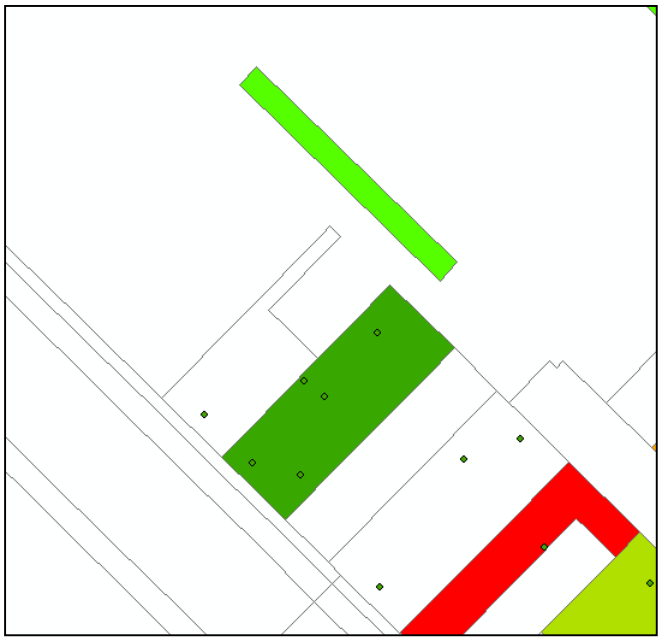


Figure 5.29: The dark green cadastral parcel, with residential and industrial function, gets the built-up culture code “Residential (agricultural”), as grassland (light green) is located near to it.

- Activity (agriculture). The reason for these wrong results is the same as for the culture code “Residential (agriculture)”; as grassland is located close to the parcels with the purpose of use “industrial function”, the parcels are given the culture code “Activity (agriculture)” instead of “Activity (industry)”.

- Residential with activity. This wrong result is caused by incorrect data in the BAG dataset. According to the BAG dataset, the cadastral parcel contains two residential objects (with residential and industrial function). However, in the real-life situation only the industrial function is present at the parcel and the parcel should, therefore, get the built-up culture code “Activity (industry)”.
- Special properties. For two cadastral parcels, the wrong result can be explained by the fact that no residential object is located on it. Both parcels contain large, secondary industrial buildings, and should have been given the built-up culture code “Activity (industry)”. The other cadastral parcel with a wrong built-up culture code “Special properties” contains a residential object with the purpose of use “other function”. However, the parcel contains a large industrial building and should, therefore, have the built-up culture code “Activity (industry)”.

Results for non-built-up culture codes

Figure 5.30 presents a map showing the cadastral parcels with a non-built-up culture code which have been selected for the quality analysis.

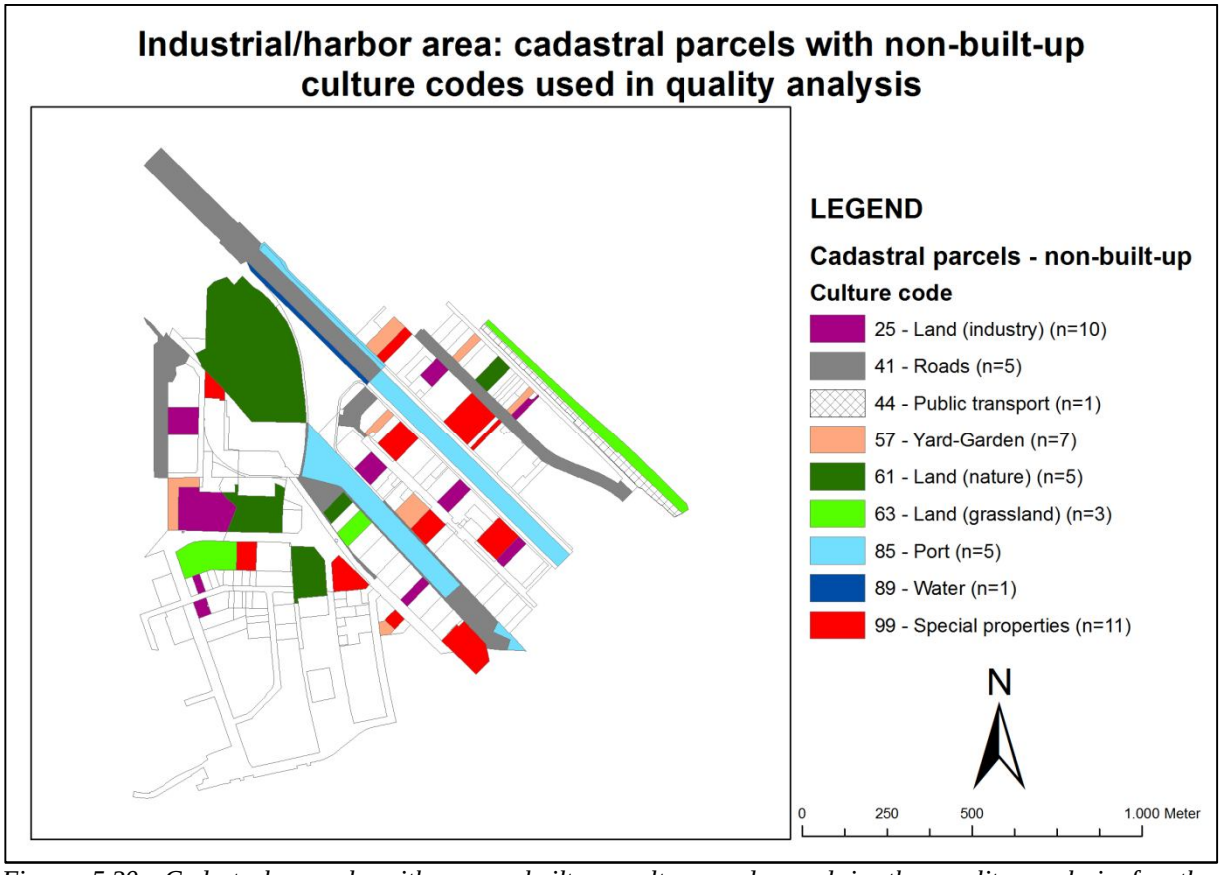


Figure 5.30: Cadastral parcels with a non-built-up culture code used in the quality analysis for the industrial/harbor test area.

Culture code	True	False	Total
25 Land (industry)	10	0	10
41 Roads	4	1	5
44 Public transport	1	0	1
57 Yard-Garden	6	1	7
61 Land (nature)	2	3	5
63 Land (grassland)	2	1	3
85 Port	5	0	5
89 Water	1	0	1
99 Special properties	11	0	11
TOTAL	***42***	***6***	***48***

Table 5.12: Results of the quality analysis for non-built-up culture codes in the industrial/harbor test area.

Table 5.12 and appendix G present the results of the quality analysis for the non-built-up culture codes in the industrial/harbor test area. The results are very satisfying, except for the non-built-up culture code "Land (nature)". For all three cadastral parcels with the wrong non-built-up culture code "Land (nature)", the wrong result can be explained by the fact that the cadastral parcel has been given the built-up culture code "Special properties" and that the cadastral parcel is partly covered by the TOP10NL terrain type "forest" (see figure 5.31 for an example). However, as the cadastral parcel should have received the built-up culture code "Activity (industry)", the non-built-up culture code should have been "Land (industry)".

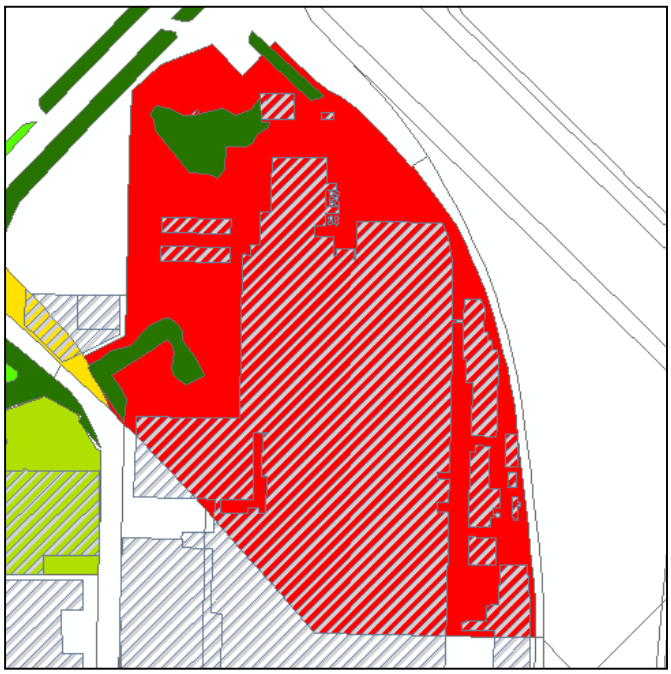


Figure 5.31: A cadastral parcel (red) covered by industrial buildings (striped) and the TOP10NL terrain type "forest" (dark green).

5.5 – Discussion

In this chapter, the decision rules and decision tree presented in the previous chapter have been tested on the cadastral parcels in four test areas in: the rural area, city centre, a living area and the industrial/harbor area of the municipality of Enschede. Three main parts of the testing can be distinguished:

1. Defining whether the cadastral parcel should get a built-up or non-built-up culture code (or both);
2. Calculating the built-up culture codes;
3. Calculating the non-built-up culture codes.

After these calculations, the test results have been tested by means of a quality analysis. This quality analysis has been performed for a number of selected cadastral parcels in each test area and for both the built-up and non-built-up culture codes.

In general, the results of this quality analysis are satisfying. A number of wrong results for the culture codes in the city centre test area can be explained by a mistake made by the author while calculating the culture codes. Re-calculating the culture codes for these particular cadastral parcels gives the right culture code. Therefore, these wrong results are not caused by the method applied, and should not be taken into account when judging the quality of the method proposed in this document.

Table 5.13 presents the number of cadastral parcels selected for the quality analysis on the calculated built-up and non-built-up culture codes. It also presents the number and percentage of right results. Taking into account all selected cadastral parcels, 80,3% of the calculated culture codes are right. For the built-up culture codes 78,2% is right and for the non-built-up culture codes 82,3%. From these numbers, one can conclude that the quality of the method proposed in this document is quite satisfying and that there is not much difference in quality between the calculated built-up and non-built-up culture codes.

Culture codes	Selected	Right	Perc. right
Built-up	170	133	78,2
Non-built-up	181	149	82,3
TOTAL	***351***	***282***	***80,3***

Table 5.13: Number and percentage of right results for built-up and non-built-up culture codes.

The most important reasons for wrong results are the following:

1. Information on the presence of catering (such as bars and restaurants) and religious buildings (such as churches) is missing. Especially in the city centre area, this has caused a number of wrong results.
2. Not all BAG buildings have a residential object. If no buildings with a residential object are located on a cadastral parcel, and if the parcel should get a built-up culture code, the parcel is given the built-up culture code “Special properties”.
3. If a BAG building with residential object is located on more than one cadastral parcel, those cadastral parcels not containing the residential object do often not get the right culture code.
4. Priority is given to parking and roads. As the percentage of the parcel covered by these land use types is not taken into account, a number of parcels dominated by another TOP10NL terrain type (such as grassland) are still given the non-built-up culture code “Parking” or “Roads”.
5. Missing or incorrect data in the source datasets (especially for the residential objects and the purpose of use in the BAG dataset).
6. The distance rule for the distinction between agricultural and industrial use of a cadastral parcel. If a residential object is within a distance of a TOP10NL terrain type that could refer to agricultural use (such as grassland and cropland), then the parcel containing this residential object is given an agriculture-related built-up culture code instead of an industry-related culture code. The same counts for the agriculture- and industry-related non-built-up culture code.

Chapter 6 – Conclusions and recommendations

In the previous chapters a production method for the culture codes dataset has been developed, presented and tested. These chapters have focused on:

1. the decision which candidate sources to use for the production of the culture codes (chapter 3);
2. formulating decision rules and integrating them into a decision tree, taking into account the possibilities and problems when using the external datasets (chapter 4);
3. a proof of concept in which culture codes have been calculated for the cadastral parcels in four case studies, and in which the results of these case studies have been tested in a quality analysis (chapter 5).

With the information and experiences obtained during the research steps carried out in these previous chapters, it is now time to get back to the research questions (see section 1.4). In this chapter answers to the main research questions are formulated (section 6.1). Furthermore, recommendations are presented for future steps concerning the production of the culture code dataset (6.2). These recommendations can be seen as the answer to the third sub-research question: what are the follow-up steps to be taken in order to be able to (automatically) produce the culture codes dataset?

6.1 – Conclusions

This section aims at formulating an answer to the first two sub-research question of this thesis. It does so by answering the following sub-questions (presented in section 1.4):

- Which source datasets could be used for an automated production of the culture code dataset? (section 6.1.1)
- To what extent is it semantically possible to use these source datasets for an automated production of the culture code dataset? (section 6.1.2)
- To what extent is it geometrically possible to use these source datasets for an automated production of the culture code dataset? (section 6.1.3)

Then, in section 6.1.4, an answer is given to the main question of this thesis.

6.1.1 Which source datasets could be used for an automated production of the culture code dataset?

After describing and analyzing the candidate source datasets, it has been decided to use the LKI, AKR, BAG and TOP10NL datasets for the production of the culture code datasets in this thesis. The contents, quality and accessibility of the candidate source datasets have been the selection criteria. For each selected dataset the reasons for using this dataset is presented:

- LKI. The LKI dataset (produced and maintained by Kadaster) is the only official dataset containing spatial data about the location, shape and size of cadastral parcels.
- AKR. The AKR dataset, also produced and maintained by Kadaster, is the administrative register of cadastral information. One of the attributes contained by the dataset is the location of apartments. As this information is not available in other candidate source datasets, it is decided to use the AKR dataset for this purpose.

- BAG. The BAG dataset is part of the Dutch system of key registers, which makes it subject to strict quality and up-to-dateness checks. The dataset contains very useful information on the location of buildings and on the usage of these buildings. As a result, the dataset is used as the primary source for the built-up culture codes to be produced. The BAG dataset is produced by municipalities, but is stored in a central database at Kadaster. Kadaster can use the spatial data stored in this database without payments involved. Finally, the BAG dataset contains historical data (going back to the moment the dataset was created, a few years ago), which makes it possible to calculate culture codes for a moment in the past.
- TOP10NL. The TOP10NL dataset has been mainly selected for its wide range of useful classes and attributes for the production of the culture codes, visualized by the visual mappings in section 3.3. The dataset contains spatial data about roads, water objects, railroads, buildings, cultural landscape objects, functional areas and terrain classes. Furthermore, the TOP10NL dataset is part of the key register BRT. As a result, the TOP10NL has a guaranteed basic level of quality. As the TOP10NL dataset is produced and maintained by Kadaster, the dataset can be used without payments involved. Finally, like the BAG dataset, the TOP10NL contains historical data.

Next to the TOP10NL dataset, the BRP and LGN datasets offer information on the agricultural and nature terrain classes. However, as the BRP dataset is not available at the moment (because of a transition in the production method) this dataset is not used in this thesis. Still, the BRP dataset could be an interesting dataset in the future, especially because of its up-to-dateness. The BRP dataset is updated every year, the TOP10NL dataset 'only' every two years.

The LGN dataset is not chosen for three main reasons: 1) as it is a grid dataset, the positional accuracy is not very good, 2) the quality and up-to-dateness are lower than those of the BRP and TOP10NL dataset, and 3) obtaining the complete LGN dataset would cost at least €63.000, while the TOP10NL dataset can be used without payment.

6.1.2 To what extent is it semantically possible to use these source datasets for an automated production of the culture code dataset?

Based on the conceptual/visual mappings between the source datasets and the culture codes (chapter 3), a list of decision rules has been formulated. This list, containing seventy-one decision rules, shows how the data from the source datasets can be used to create the built-up and non-built-up culture codes. Furthermore, the decision rules have been integrated into a decision tree, in order to visualize how the information from these source datasets is used for the production of the culture codes in this thesis.

As the source datasets have also been selected on their contents – next to their contents, quality and accessibility –, it was expected that most culture codes can be produced by making use of these datasets. However, for a small number of culture codes the datasets do not offer useful information. This applies to the culture codes "Pipes – Tubes" and "Activity (glasshouse)". These culture codes can, therefore, not be produced by making use of the selected datasets. Furthermore, a number of built-up culture codes relies on optional attributes from the TOP10NL dataset: Storeroom – shelter (garage-shed), Activity (utility), Activity (catering), Parking, Defense, Public transport, Air traffic, Culture, Police – Fire service, Religion, Funeral service and Port. The non-built-up culture code "Structures – Water works" also relies on optional information. These culture codes might, as a result, have a lower completeness.

6.1.3 To what extent is it geometrically possible to use these source datasets for an automated production of the culture code dataset?

The source datasets contain a combination of point, line and polygon geometries. Polygon geometries make it possible to calculate percentages of overlap between cadastral parcels and other spatial data, e.g. BAG buildings. This kind of calculations is not possible with points and lines. As a result, it is necessary to work with priorities. For example, if a railroad section (with a line geometry) is located on a cadastral parcel, this land use type should get priority over a terrain class grassland (with a polygon geometry), instead of calculating the culture code by making use of the percentages of overlap. In some cases it is possible to perform a spatial join between information sources with point and polygon geometries. For example, the purposes of use of the BAG residential objects (points) have been spatially joined to the cadastral parcels (polygons).

Percentages of overlap are to be used in case different terrain classes are located on one cadastral parcel. For example, if 15% of the parcel is covered by water and 85% by grassland, then the cadastral parcel should get the culture code "Land (grassland)". Such a method is, however, not possible when dealing with point objects; for example, if more than one purpose of use type is located on a cadastral parcel. Some culture codes refer to a mixed land use (such as "Residential with activity"), otherwise the cadastral parcel is given the culture code "Special properties".

Another issue arises when deciding whether a cadastral parcel should get a built-up or non-built-up culture code or a combination. In some cases BAG buildings cover a very small part of a cadastral parcel as they just cross a parcel boundary. In such cases the cadastral parcel covered by only a very small part of the building should not get a built-up culture code. As a result, it has been decided to calculate the percentage of overlap between cadastral parcels and BAG buildings. When less than 2% of a parcel is covered by BAG buildings, that parcel does not get a built-up culture code.

The seventy-one decision rules and the decision tree have been proposed in order to be able to produce the culture codes by making use of the attributes from the source datasets. However, they should also help to overcome the geometrical issues. Therefore, priorities are found in the decision tree, as a certain order in the steps is proposed.

The proof of concept has shown that the decision rules and decision tree give quite satisfying results with 80,3% of the calculated culture codes being the right code. As a result, it can be concluded that the geometrical issues can, to a large extent, be overcome by the production method proposed in this thesis.

6.1.4 General conclusion

In this section an answer is formulated to the main question of this thesis:

> *To what extent is it possible to automatically produce a culture code (land use) dataset by making use of external source datasets, when taking into account their quality as well as their semantic, geometric and temporal characteristics?*

As has been described in the previous two sections, in this thesis a number of source datasets for the production have been selected by taking into account their quality and their semantic, geometric and temporal characteristics. Four source datasets have been selected (LKI, AKR, BAG and TOP10NL), with the decision using these datasets being based on their characteristics. Afterwards, the visual mappings between the classes and attributes of these datasets and the culture codes have been translated into decision rules. These decision rules have been integrated into a decision tree which has been tested in the proof of concept.

The proof of concept shows quite satisfying results, as 80,3% of the calculated culture codes proved to be right in the quality analysis. This means that, by making use of a limited number of source datasets, it is to a very large extent possible to produce the culture codes. However, a number of limitations of the proposed method have been identified:

1. The large number of features used in the calculations can lead to serious performance problems. In this thesis, performance problems occurred when trying to calculate the culture codes for the municipality of Enschede. Therefore, it was chosen to use four smaller case studies for the proof of concept.
2. Two culture codes cannot be produced when using the selected source datasets: "Pipes – Tubes" and "Activity (glasshouse)".
3. A number of culture codes relies completely on optional attributes from the TOP10NL object class "type of building". The completeness of these culture codes could therefore be an issue.
4. Information on the location of catering and religious buildings is missing in the datasets.
5. A large number of built-up cadastral parcels have been given the culture code "Special properties", as the buildings located on these parcels do not contain a residential object.
6. When a BAG building is located on more than one parcel and not all of these parcels contain a residential object, some of the parcels get the culture code "Spatial properties".
7. Because of the priority given to parking and roads, some cadastral parcels dominated by for example grassland get the culture code "Parking" or "Roads" instead of "Land (grassland)".
8. In some cases cadastral parcels containing industrial activity are given agriculture-related culture codes. As the BAG dataset does not make a difference between industry and agriculture, distances between residential objects and agriculture-related terrain classes are used to define whether a cadastral parcel should get an industry- or agriculture-related culture code. This does not always give satisfying results, especially in the industrial/harbor area.

6.2 – Recommendations

This section presents recommendations for further research and steps to be performed in order to introduce an automated production of the culture codes dataset, based on the experiences from the research carried out in this thesis. A general recommendation is that using external source datasets for the production of the culture codes should be seen as a serious option, as the proof of concept in this thesis has given quite satisfying results. However, a number of issues needs further attention. These issues can be clustered into three main areas of attention. These three clusters and the recommendations related to them are the following:

1. Additional geographical information. The analysis of the candidate source datasets and the proof of concept have shown that additional geographical information might become available in the future and is needed to further improve the quality of the calculated culture codes. Recommendations for this cluster are presented in table 6.1.

Recommendation	Description
The developments concerning the BGT dataset should be followed carefully	The BGT, unlike the current GBKN dataset, contains polygon objects. This would make the dataset useful for making calculations on the percentage of overlap. Furthermore, the up-to-dateness and quality of the dataset will be improved. Finally, the BGT will have more detail than the TOP10NL, which could give better results in the culture code calculations.
When again available, the usefulness of the BRP dataset should be investigated	At the moment, this dataset is not available, but because of its detailed information on the usage of the agricultural parcels and its up-to-dateness it might be useful additional information for the production of the culture codes.
Additional information on the location of catering buildings should be used	This information is missing in the selected source datasets, but might be available in other geographical datasets.
Additional information on the location of religious buildings should be used	Also for religious buildings information is missing to a large extent. Therefore, it might be necessary to use another geographical dataset.
For those culture codes relying on optional attributes from the TOP10NL dataset, additional information should be used	As these attributes are optional, the completeness of the culture codes relying on them cannot be guaranteed. By using additional geographical datasets, it might be possible to increase the completeness.

Table 6.1: Recommendations related to additional geographical information.

2. Next steps concerning the production method. Next to adding new information, the production method proposed in this thesis might get further attention. It might be necessary to change the method to get better results. Recommendations related to this issues are presented in table 6.2.

Recommendation	Description
The distances to agricultural terrain classes used to define a difference between agriculture- and industry-related culture codes should be investigated	In order to make a difference between agricultural- and industry-related culture codes distances between residential objects and agricultural terrain classes have been used (more or less than 50m). The proof of concept has shown that this rule can lead to wrong results, especially in the industrial/ harbor area. It might therefore be better to work with another distance or to use another method for making the distinction between agriculture- and industry-related culture codes. This needs further investigation.
The priority given to, for example, parking and roads needs further investigation	Because of the priority given to parking and roads, some cadastral parcels dominated by other land use types (e.g. grassland) were still given the culture code “Parking” or “Roads”. It should be investigated whether using percentages of overlap with the cadastral parcels (such as for water, grassland, etc.) instead of priorities would give better results for the culture codes.
The impact of using a spatial join between residential objects and buildings (instead of cadastral parcels) should be investigated	In the current production method, a spatial join between cadastral parcels and residential objects (with purpose of use) is used. In some cases, when a building is located on more than one cadastral parcel, this has lead to wrong results for the parcels without residential object. It should be investigated what would be the results when using a spatial join between buildings and residential objects.

Table 6.2: Recommendations related to the production method.

3. Issues requiring special attention. Finally, a number of issues require special attention. The recommendations which belong to this cluster are presented in table 6.3.
In table 6.3 a distinction is made between recommendations about the general culture code project and those which are important when it is decided to implement an automated way of producing the culture codes based on the method proposed in this thesis.

Recommendation	Description
GENERAL CULTURE CODE PROJECT	
It should be decided which culture codes are really needed for fulfilling the needs of the users	This is carried out by the Kadaster research group for the culture code production. It could be that some culture codes are not needed anymore. The production method should be adjusted to the user needs.
A business case should be carried out in order to define the best way of producing the culture codes dataset	An automated production of the culture codes dataset by making use of external source datasets is just one possible method for the production of the dataset. It should be investigated if this is the most efficient and effective way of producing. Maybe other production ways give better results or maybe they are easier to implement.
It should be considered to start an investigation on the option of using multiple culture codes per cadastral parcel.	In many cases, cadastral parcels contain more than one land use type. At the moment, it is only possible to have multiple culture codes when these are a built-up and a non-built-up culture code. The other types of land use contained by the cadastral parcel are, therefore, not represented by the culture code dataset. To make some of this 'lost' spatial information part of the dataset, it might be an option to have multiple culture codes per cadastral parcel.
It should be considered to start an investigation on the option of creating sub-parcels based on the culture codes.	Next to the option of using multiple culture codes per cadastral parcel, it is an option to create sub-parcels based on the culture codes. Again, this would mean that less data on land use is 'lost', as the sub-parcels would (ideally) contain only one land use type. This counts more for the non-built-up culture codes than for the built-up culture codes. Buildings often have two purposes of use. The terrain types (TOP10NL) are partitioning and not overlapping. However, the object classes with a point or line geometry are overlapping the terrain types. So, even this option would not prevent the culture codes dataset from 'losing' data on land use.
RELATED TO THE METHOD PROPOSED IN THIS THESIS	
It should be investigated how the performance issues can be overcome	Because of the large amount of spatial data, performance problems arose when calculating the culture codes for the whole municipality of Enschede. As the culture codes should be calculated for the Netherlands, the current method might cause performance problems. Therefore, it is advised to search for a way to overcome these problems. Things that could be investigated, are: hardware and cutting up the country in smaller areas.

A suitable tool has to be chosen for the automated production	A tool has to be chosen for an automated mapping procedure. It depends on the schema modeling language used, which tool should be chosen. Examples of tools available are HALE (Humboldt Alignment Editor) (Humboldt Community 2011) and FME Desktop (Safe Software 2011b).

Table 6.3: Recommendations related to issues requiring special attention.

Bibliography

Alexe, B., L. Chiticariu, R. J. Miller, and W.-C. Tan. 2008. Muse: mapping understand and deSign by example. In *Proceedings of the IEEE 24th International Conference on Data Engineering*. Cancun, Mexico, 7-12 April 2008.

Alterra. 2011. *Landelijk Grondgebruik Nederland* [cited 15 March 2011]. Available from http://www.alterra.wur.nl/NL/onderzoek/Werkveld+Geoinformatie /Producten_Services/LGN/.

Ambler, S. W. 2010. *Introduction to UML 2 class diagrams* [cited 6 November 2011]. Available from http://www.agilemodeling.com/artifacts/classDiagram.htm.

Bakker, N., B. Bruns, and M. Storm. 2005. *Gegevensmodel TOP10NL* [cited 13 January 2011]. Available from http://www.kadaster.nl/top10nl/gegevensmodel _top10nl_2%203.pdf.

Batini, C., M. Lenzerini, and S. B. Navathe. 1986. A comparative analysis of methodologies for database schema integration. *ACM Computing Surveys* 18 (4): pp. 323-364.

Brassel, K., F. Bucher, E.-M. Stephan, and A. Vckovski. 1995. Completeness. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 81-108. Oxford: Elsevier Science Ltd.

Cambridge University Press. 2011. *Cambridge Dictionaries Online* [cited 10 November 2011]. Available from: http://dictionary.cambridge.org/.

Clarke, D. G., and D. M. Clark. 1995. Lineage. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 13-30. Oxford: Elsevier Science Ltd.

Crompvoets, J. 2009. Introduction to spatial data infrastructures and INSPIRE. Wageningen, the Netherlands, 18 March 2009 [lecture].

de Man, W. H. E. 2007. Are spatial data infrastructures special? In *Research and theory in advancing spatial data infrastructure concepts*, ed. H. Onsrud, pp. 33-54. Redlands, California, USA: ESRI Press.

Devogele, T., C. Parent, and S. Spaccapietra. 1998. On spatial database integration. *International Journal of Geographical Information Science* 12 (4): pp. 335-352.

Dijkstra, H. 2005. *Monitoring en evaluatie Agenda Vitaal Platteland. Inventarisatie aanbod monitoringsystemen* [cited 10 November 2011]. Available from: http://library.wur.nl/ way/bestanden/clc/1794183.pdf.

Drummond, J. 1995. Positional accuracy. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 31-58. Oxford: Elsevier Science Ltd.

E-overheid. 2009. *Betere dienstverlening door e-overheid en NUP* [cited 27 October 2009]. Available from http://www.e-overheid.nl/home.html.

Ellis, E. 2010. *Land-use and land-cover change* [cited 27 November 2010]. Available from http://www.eoearth.org/article/Land-use_and_land-cover_change.

ESRI. 2009. *ArcGIS Desktop 9.3 Help* [cited 3 October 2011]. Available from http://webhelp.esri.com/arcgiSDEsktop/9.3/index.cfm?TopicName=welcome.

Fagin, R., L. M. Haas, M. Hernández, R. J. Miller, L. Popa, and Y. Velegrakis. 2009. Clio: schema mapping creation and data exchange. *Lecture Notes in Computer Science* 5600: pp. 198-236.

Feeney, M., and I. P. Williamson. 2000. Researching frameworks for evolving Spatial Data Infrastructure. In *SIRC 2000 - The 12th Annual Colloquium of the Spatial Information Research Centre*. Dunedin, New Zealand.

Foerster, T., L. Lehto, T. Sarjakoski, L. T. Sarjakoski, and J. E. Stoter. 2010. Map generalization and schema transformation of geospatial data combined in a Web Service context. *Computers, Environment and Urban Systems* 34 (1): pp. 79-88.

Fuld, M., and M. Rietdijk. 2004. Met zijn allen recht op het doel af. De strategie voor de implementatie van de basisregistraties voor adressen en gebouwen. Den Haag: Ministerie van VROM.
Geonovum. 2009. *ISO 19115 Metadata standaard* [cited 10 November 2011]. Available from: http://wiki.geonovum.nl/index.php/1.5.1_ISO_19115_Metadata_standaard.
GIS Service Agro.nl. 2011. *BRPGewas2009* [cited 21 May 2011]. Available from http://gisserver1.agro.nl/arcgis/rest/services/BRPGewas2009/MapServer.
Goodchild, M. F. 1995. Attribute accuracy. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 59-80. Oxford: Elsevier Science Ltd.
———. 2000. Communicating geographic information in a digital age. *Annals of the Association of American Geographers* 90 (2): pp. 344-355.
Guptill, S. C. 1995. Temporal information. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 153-166. Oxford: Elsevier Science Ltd.
Hendriks, W., and J. Dijk. 2011. Interview about the BRP dataset. Assen, the Netherlands, 27 April 2011 [personal communication].
Heywood, I., S. Cornelius, and S. Carver. 2006. *An introduction to Geographical Information Systems*. Third ed. Harlow, England: Pearson Education Limited.
Humboldt Community. 2011. *Humboldt Alignment Editor - HALE* [cited 23 January 2011]. Available from http://community.esdi-humboldt.eu/wiki/hale.
ISO: International Organization for Standardization. 2011. *ISO 19115:2003 – Geographic Information – Metadata* [cited 10 November 2011]. Available from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020.
ISO TC 211/SC. 2010. ISO/DIS 19152: Geographic information - Land Administration Domain Model (LADM) [draft version, 20 July 2010].
Kadaster. 2004a. AKR kultuurindicaties transactie 04. Apeldoorn: Kadaster.
———. 2004b. Indikken cultuurcodes algemeen. Apeldoorn: Kadaster.
———. 2007. *Catalogus Basisregistratie Topografie. Versie 2.0* [cited 17 January 2011]. Available from http://www.e-overheid.nl/e-overheid-2.0/live/binaries/stelselhandboek/gegevens/catalogusbrt2_0.pdf.
———. 2010. *Kadata - Veel gestelde vragen* [cited 12 October 2010]. Available from https://kadata.kadaster.nl/asp/faq/veelgesteldevragen1.asp.
———. 2011a *Handleiding massale output AKR (GeAutomatiseerde Kadastrale Registratie)* [cited 6 November 2011]. Available from www.kadaster.nl/codestelsel/MO_handleiding.doc.
———. 2011b. *TOP10NL* [cited 13 January 2011]. Available from http://www.kadaster.nl/top10nl/.
Kainz, W. 1995. Logical consistency. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 109-138. Oxford: Elsevier Science Ltd.
Kooij, F. 2011. Interview about the BAG dataset. Apeldoorn, the Netherlands, 21 April 2011 [personal communication].
Kraak, M.-J., and F. Ormeling. 2003. *Cartography: visualization of geospatial data*. Second ed. Harlow, England: Pearson Education Limited.
LSV GBKN. 2009. *GBKN Productinformatie* [cited 13 January 2011]. Available from http://www.gbkn.nl/nieuwesite/html/productinformatie.html#top.
Maessen, B. 2011a. Interview about spatial data quality elements and GBKN/BGT. Apeldoorn, The Netherlands, 30 September 2011 [personal communication].
———. 2011b. Interview about the GBKN and BGT datasets. Apeldoorn, the Netherlands, 30 March 2011 [personal communication].
Mayhew, S. 2004. *A dictionary of geography. Third edition*. New York, USA: Oxford University Press Inc.

McLaughlin, J., and S. E. Nichols. 1989. Resource management: the land administration and cadastral systems component. *Surveying and Mapping* 49 (2): pp. 77-86.
Miller, R. 2003. *Practical UML: A hands-on introduction for developers* [cited 26 October 2010]. Available from http://edn.embarcadero.com/article/31863.
Ministerie van LNV. 2011. *Meldingen WBP: Basisregistratie percelen* [cited 17 January 2011]. Available from http://www2.minlnv.nl/cgi-bin/xmlparser.pl?config=/lnv/algemeen/wbp&file=wbp_meldingen_dr/primair/basisregistratie_percelen.xml#basisgegevens.
Ministerie van VROM. 2009a. *Catalogus basisregistraties adressen en gebouwen. Versie 2009* [cited 17 January 2011]. Available from http://www.kadaster.nl/BAG/docs/catalogus_grondslagenBAG.pdf.
———. 2009b. *Objectenhandboek basisregistraties adressen en gebouwen. Versie 2009* [cited 17 January 2011]. Available from http://www.kadaster.nl/BAG/docs/objectenhandboek_2009.pdf.
———. 2009c. *Processenhandboek basisregistraties adressen en gebouwen. Versie 2009* [cited 17 January 2011]. Available from http://www.kadaster.nl/BAG/docs/processenhandboek.pdf.
Morrison, J. L. 1995. Spatial data quality. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 1-12. Oxford: Elsevier Science Ltd.
NEN: Nederlands Normalisatie-Instituut. 2005. *NEN 3610:2005. Basismodel geo-informatie - termen, definities, relaties en algemene regels voor de uitwisseling van informatie over aan het aardoppervlak gerelateerde ruimtelijke objecten*. Delft: NEN.
———. 2011. *NEN3610:2010 Ontw. Basismodel geo-informatie - Termen, definities, relaties en algemene regels voor de uitwisseling van informatie over aan de aarde gerelateerde ruimtelijke objecten*. Delft: Nederlands Normalisatie-Instituut.
Palmer, D., and J. McLaughlin. 1996. Integrated land administration: institutional and technical challenges. *A special issue of the ITC Journal for the Habitat II Conference* 1996 (1).
Peersmann, M. 2009. "Richtlijn voor regionale GBKN stichtingen": van GBKN naar BGT in Gr/Dr/Ov. Emmen, The Netherlands, 19 November 2009 [presentation].
Programma BGT, A. I. S. 2010. Informatiemodel Basisregistratie Grootschalige Topografie (BGT). Gegevenscatalogus, versie 0.9_7. Den Haag: Ministerie van Infrastructuur en Milieu, Portefeuille Ruimte.
Rajabifard, A., and I. P. Williamson. 2001. Spatial data infrastructures: concept, SDI hierarchy and future directions. In *Geomatics 80*. Tehran, Iran.
Rajabifard, A., I. P. Williamson, P. Holland, and G. Johnstone. 2000. From local to global SDI initiatives: a pyramid of building blocks. In *4th Global Spatial Data Infrastructure Conference*. Cape Town, South Africa.
Safe Software. 2011a. *Clipper* [cited 6 November 2011]. Available from http://docs.safe.com/fme/2009/html/Transformers/content/transformers/clipper.htm.
———. 2011b. *FME Desktop* [cited 7 November 2011]. Available from http://www.safe.com/fme/fme-technology/fme-desktop/overview/.
Salgé, F. 1995. Semantic accuracy. In *Elements of spatial data quality*, eds. S. C. Guptill and J. L. Morrison, pp. 139-152. Oxford: Elsevier Science Ltd.
Schwering, A. 2008. Approaches to semantic similarity measurement for geo-spatial data: a survey. *Transactions in GIS* 12 (1): pp. 5-29.
SpatialEcology.com. 2011. *Hawth's Analysis Tools for ArcGIS* [cited 25 October 2011]. Available from http://www.spatialecology.com/htools/tooldesc.php.
Stuiver, J. 2010. Lineage in GIS. Wageningen, the Netherlands, 25 March 2010 [lecture].

te Winkel, D. 2011a. Interview about spatial data quality elements and TOP10NL. Apeldoorn, the Netherlands, 21 October 2011 [personal communication].
———. 2011b. Interview about the TOP10NL and BRT datasets. Apeldoorn, the Netherlands, 29 March 2011 [personal communication].
van den Brink, T. 2010. *Productbeschrijving LKI percelenkaart* [cited 5 November 2011]. Available from www.amsterdam.nl/publish/.../productbeschrijving_lki-versie-0.6.pdf.
van Loenen, B., J. A. Zevenbergen, and J. de Jong. 2008. Geo-informatie: wat is het en wat is de juridische context? In *Recht en locatie: geo-informatie in een juridische context*, eds. J. G. L. van der Wees and J. Nouwt, pp. 11-33. Den Haag: Reed Business.
van Oort, P. A. J. 2006. Spatial data quality: from description to application, Wageningen University, the Netherlands.
Van Oosterom, P. J. M., C. Lemmen, H. Uitermark, G. Boekelo, and G. Verkuijl. 2011. *Land administration standardization with focus on surveying and spatial representations* [cited 10 November 2011]. Available from: http://www.surveysummit.com/2011-proceedings/files/land-administration-standardization.pdf.
van Rossem, R. 2009. Basisregistratie Grootschalige Topografie. Deventer, 26 November 2009 [presentation].
W3C. 2004. *Web Ontology Language (OWL)* [cited 17 December 2010]. Available from http://www.w3.org/2004/OWL/.
Worboys, M. F. 1994. A unified model for spatial and temporal information. *The Computer Journal* 37 (1): pp. 26-34.
Xu, W., A. Dilo, S. Zlatanova, and P. J. M. van Oosterom. 2008. Modelling emergency response processes: comparative study on OWL and UML. In *Information systems for crisis response and management. Proceedings of the Joint ISCRAM-CHINA and G14DM Conference*. Harbin, China, 4-6 August 2008.
Zevenbergen, J. A. 2002. *Systems of land registration - aspects and effects*. Delft, The Netherlands: NCG, Netherlands Geodetic Commission.
———. 2004. A systems approach to land registration and cadastre. *Nordic Journal of Surveying and Real Estate Research* 1 (1): pp. 11-24.

Appendix A – List of culture codes

List with the culture codes (in Dutch and translated to English). The culture codes in red are not used anymore.

Kultuur kode/ Culture code	Omschrijving	Description
0	initieel	Initial
11	WONEN	Residential
12	WONEN (APPARTEMENT)	Residential (apartment)
13	Bijzondere woonvormen	Special types of dwelling
14	WONEN (AGRARISCH)	Residential (agricultural)
15	Recreatie (tweede woning)	Recreation (second home)
16	Woonwagenkampen	Caravan sites
17	Woonbotenhavens	Houseboat marinas
18	BERGING-STALLING (GARAGE-SCHUUR)	Storeroom – shelter (garage-shed)
19	Overige woonvormen	Other types of dwelling
21	BEDRIJVIGHEID (KANTOOR)	Activity (office)
22	BEDRIJVIGHEID (INDUSTRIE)	Activity (industry)
23	BEDRIJVIGHEID (NUTSVOORZIENING)	Activity (utility)
24	Laboratoria aktiviteit (kontrole, meten onderzoek)	Laboratory activity (inspection, testing, research)
25	TERREIN (INDUSTRIE)	Land (industry)
26	Opslag	Storage
27	BEDRIJVIGHEID (DETAILHANDEL)	Activity (retail trade)
28	BEDRIJVIGHEID (HORECA)	Activity (catering)
29	BEDRIJVIGHEID (AGRARISCH)	Activity (agricultural)
31	Wonen met kantoorbedrijvigheid	Residential with office activity
32	Wonen met produktie, reparatie en onderhoud	Residential with production, repair and maintenance
33	Wonen met nutsbedrijven	Residential with utilities
34	PARKEREN	Parking
35	DEFENSIE	Defense
36	Wonen met opslag	Residential with storage
37	WONEN MET BEDRIJVIGHEID	Residential with activity
38	Wonen met horeca	Residential with catering
39	Wonen met overige bedrijvigheid	Residential with other activity
41	WEGEN	Roads
42	Semi-verharde wegen	Semi-paved roads
43	Niet verharde wegen	Unpaved roads
44	OPENBAAR VERVOER	Public transport
45	Tram- en metrowegen	Tramways and underground railways
46	LUCHTVERKEER	Air traffic
47	LEIDINGEN - BUIZEN	Pipes – Tubes
48	Buisleidingtransport	Pipeline transport
49	Elektriciteitstransport	Electricity transport
51	RECREATIE - SPORT	Recreation – Sport
52	Volkstuinen	Allotments
53	WONEN (RECREATIE)	Residential (recreation)
54	Recreatie-objekten	Recreational properties

55	PARKEN - PLANTSOENEN	Parks – public gardens
56	Bos met recreatieve hoofdfunktie	Wood – main function recreational
57	ERF - TUIN	Yard – Garden
59	Overige recreatie	Other recreation
61	TERREIN (NATUUR)	Land (nature)
62	TERREIN (AKKERBOUW)	Land (farming)
63	TERREIN (GRASLAND)	Land (grassland)
64	BEDRIJVIGHEID (KAS)	Activity (glasshouse)
65	Tuinbouw in de open grond	Horticulture in the open ground
66	TERREIN (TEELT - KWEEK)	Land (cultivation)
67	Bloembollengrond	Bulb-growing land
68	Boomkwekerijen en kerstdennenkultuur	Tree nurseries and Christmas tree cultivation
71	Gezondheid	Health
72	GEZONDHEID	Health
73	Dagverblijven	Day-care centres
74	ONDERWIJS	Education
75	Kulturele aktiviteit	Cultural activity
76	CULTUUR	Culture
77	GODSDIENST	Religion
78	POLITIE - BRANDWEER	Police – Fire service
79	JUSTITIE	Justice
81	Buitenwater	Open water
82	Waterreservoirs	Water reservoirs
83	Gracht, vaart, kanalen	Canals
84	Meren, plassen, vennen	Lakes, ponds, marshes
85	HAVEN	Port
86	Zeehavens	Seaports
87	BOUWWERKEN - WATERWERKEN	Structures – Water works
88	Rivieren	Rivers
89	WATER	Water
91	Braak terrein	Fallow land
92	TERREIN NIEUWBOUW-BEDRIJVIGHEID	Land (new construction activity)
93	TERREIN NIEUWBOUW-WONEN	Land (new construction residential)
94	UITVAART	Funeral service
95	Stortplaatsen	Dumps
96	Wrakkenopslagplaatsen	Scrap car storage sites
97	Droog natuurlijk terrein	Dry natural land
98	Nat natuurlijk terrein	Wet natural land
99	BIJZONDERE OBJECTEN	Special properties

Source: Kadaster (2010).

Appendix B – Translation table

This table gives an overview of Dutch terms translated into English. It gives only those terms for which the translation is not given in the main body of this thesis. The order is as they appear in the thesis.

Dataset	Dutch term	English translation
BAG	Panden	Buildings
	Verblijfsobjecten	Residential object
	Gebruiksdoeleind	Purpose of use
	Woonfunctie	Residential function
	Winkelfunctie	Retail trade function
	Bijeenkomstfunctie	Gathering function
	Logiesfunctie	Accommodation function
	Gezondheidszorgfunctie	Health function
	Industriefunctie	Industry function
	Sportfunctie	Sports function
	Kantoorfunctie	Office function
	Onderwijsfunctie	Education function
	Cel functie	Prison/Cells function
	Overige gebruiksfunctie	Other function
	Status	Status
	Bouw gestart	Building process started
	Bouwvergunning verleend	Building permit granted
BRP	Bouwland	Cropland
	Grasland	Grassland
	Natuurterrein	Nature area
	Braakland	Bare soil
TOP10NL	Wegdeel	Road section
	Hoofdverkeersgebruik	Main use
	Busverkeer	Bus traffic
	Ruiters	Horse riding
	Vliegverkeer	Air traffic
	Parkeren	Parking
	Parkeren: P+R parkeerplaats	Park + ride
	Parkeren: carpoolplaats	Carpool site
	Gemengd verkeer	Mixed traffic
	Snelverkeer	Fast traffic
	Spoorbaandeel	Railway section
	Type spoorbaan	Type of railway
	Trein	Train
	Tram	Tram
	Metro	Metro/Subway
	Gemengd	Mixed
	Waterdeel	Water object
	Type water	Type of water
	Waterloop	Water stream
	Meer, plas, ven, vijver	Lake, fen, etc.
	Greppel, droge sloot	Ditch

	Zee	Sea
	Droogvallend	Tidal
	Bron, wel	Well
	Fysiek voorkomen	Physical appearance
	In sluis	In a sluice
	Functie	Function
	Haven	Port
	Natuurbad	Natural pool
	Vistrap	Fish ladder
	Waterzuivering	Water treatment
	Zwembad	Swimming pool
	Gebouw	Building
	Type gebouw	Type of building
	Bezoekerscentrum	Visitor centre
	Crematorium	Crematory
	Dok	Dock
	Elektriciteitscentrale	Powerplant
	Fort	Plant/Factory
	Gemaal	Pumping station
	Gevangenis	Prison
	Hotel	Hotel
	Huizenblok	Housing block
	Kapel	Chapel
	Kerk	Church
	Kerncentrale, kernreactor	Nuclear powerplant
	Klooster, abdij	Monastery
	Kliniek, inrichting, sanatorium	Clinic
	Kunstijsbaan	Artificial ice track
	Manege	Horse riding school
	Metrostation	Metro station
	Militair gebouw	Military building
	Motel	Motel
	Museum	Museum
	Parkeerdak, parkeerdek, parkeergarage	Parking garage
	Politiebureau	Police office
	Pompstation	Petrol station
	Psychiatrisch ziekenhuis, psychiatrisch centrum	Psychiatric clinic
	Recreatiecentrum	Recreation centre
	Reddingboothuisje	Lifeboat building
	Religieus gebouw	Religious building
	Remise	Coach-house
	School	School
	Sporthal	Sports building
	Stadion	Stadium
	Treinstation	Train station
	Universiteit	University
	Verkeerstoren	Traffic tower
	Wegrestaurant	Road restaurant
	Werf	Shipyard
	Ziekenhuis	Hospital

	Zwembad	Swimming pool
	Terrein	Terrain
	Type landgebruik	Land use type
	Aanlegsteiger	Landing-stage/Pier
	Akkerland	Cropland
	Boomgaard, boomkwekerij	Tree nursery
	Bos	Forest
	Fruitkwekerij	Fruit farm
	Grasland	Grassland
	Heide	Heath/Moor
	Inrichtingselement	Cultural landscape object
	Type inrichtingselement	Type of object
	Aanlegsteiger	Landing-stage/Pier
	Gaswinning	Gas extraction
	Gedenkteken, monument	Memorial site/Monument
	Gemaal	Pumping station
	Helikopterlandingsplatform	Helicopter platform
	Hoogspanningsleiding	Power line
	Kapel	Chapel
	Schietbaan	Shooting range
	Sluisdier	Sluice door
	Stormvloedkering	Water barrier
	Station	Station
	Stuw	Dam
	Registratief gebied	Administrative area
	Type registratief gebied	Type of administrative area
	Nationaal park	National park
	Geografisch gebied	Geographical area
	Type geografisch gebied	Type of geographical area
	Bosgebied	Forest area
	Duingebied	Dune area
	Heidegebied	Moor/Heath area
	Meer, plas, ven, vijver	Lake, fen, etc.
	Wad	Mudflat
	Zee	Sea
	Zeegat, zeearm	Sea arm
	Functioneel gebied	Functional area
	Type functioneel gebied	Type of functional area
	Arboretum	Arboretum
	Bedrijventerrein	Business area
	Begraafplaats	Cemetery
	Bungalowpark	Bungalow park
	Camping, kampeerterrein	Camping place
	Caravanpark	Caravan park
	Circuit	Circuit
	Crossbaan	Crossing track
	Dierentuin, safaripark	Zoo
	Golfterrein	Golf course
	Grafheuvel	Burial mound
	Haven	Port
	Helikopterlandingsterrein	Helicopter landing site

	Jachthaven	Yacht harbor
	Kartingbaan	Karts track
	Militair oefengebied, schietterrein	Military area
	Mijn	Mine
	Mosselbank	Mussel bed
	Natuurgebied, natuurreservaat	Nature area
	Openluchtmuseum	Open air museum
	Openluchttheater	Open air theatre
	Park	Park
	Recreatiegebied	Recreational area
	Renbaan	Hippodrome
	Skibaan	Ski slope
	Sluizencomplex	Sluice
	Sportterrein, sportcomplex	Sports site
	Stortplaats	Dump
	Tennispark	Tennis course
	Tuincentrum	Garden centre
	Vliegveld, luchthaven	Airport
	Volkstuinen	Allotment
	Werf	Shipyard
	IJsbaan	Ice track
	Zuiveringsinstallatie	Water treatment
GBKN/BGT	Wegdeel	Road section
	Type weg	Type of road
	Baan voor vliegverkeer	Starting lane for air traffic
	Autosnelweg	Highway
	Hoofdweg	Main road
	Regionale weg	Regional road
	Lokale weg	Local road
	Straat	Street
	Pad	Path
	Waterdeel	Water object
	Type water	Type of water object
	Zee	Sea
	Waterloop	Water stream
	Watervlakte	Water surface
	Kunstwerk	Engineering object
	Type kunstwerk	Type of engineering object
	Brug	Bridge
	Gemaal	Pumping station
	Perron	Platform (public transport)
	Sluis	Sluice
	Strekdam	Dam
	Tunnel	Tunnel
	Windturbine	Wind turbine
	Stuw	Weir
	Trap	Stairs
	Scheiding	Separation element
	Type scheiding	Type of separation element
	Muur	Wall
	Kademuur	Quay/Pier wall

	Geluidsscherm	Noise screen
	Damwand	Dam wall
	Walbescherming	Shore protection
	Heg	Hedge
	Hek	Fence
	Overig bouwwerk	Other built elements
	Type overig bouwwerk	Type of built element
	Open loods	Open stall
	Bezinkbak	Settling tank
	Lage trafo	Low transformer
	Bassin	Basin
	Silo	Silo
	Steiger	Landing-stage/Pier
AKR	Appartementsrechten	Apartment rights

Appendix C – List of definitions for culture codes

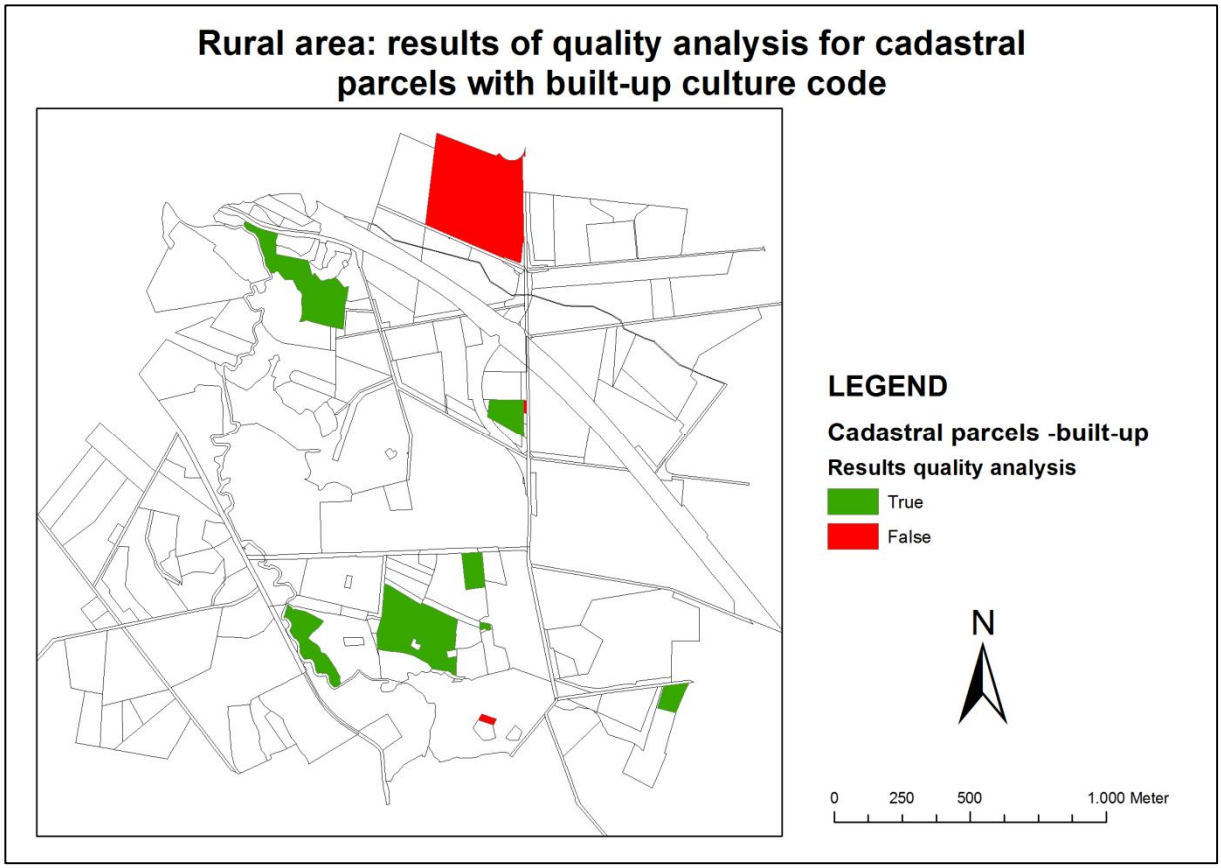
This table provides a definition for each culture code currently used at the institution Kadaster (the black culture codes in Appendix A). The list provides, when available, definitions from the online dictionary *Cambridge Dictionaries Online* (Cambridge University Press 2011). These definitions are put between quotation marks. Furthermore, when needed, information about the content of this culture code is added by the author.

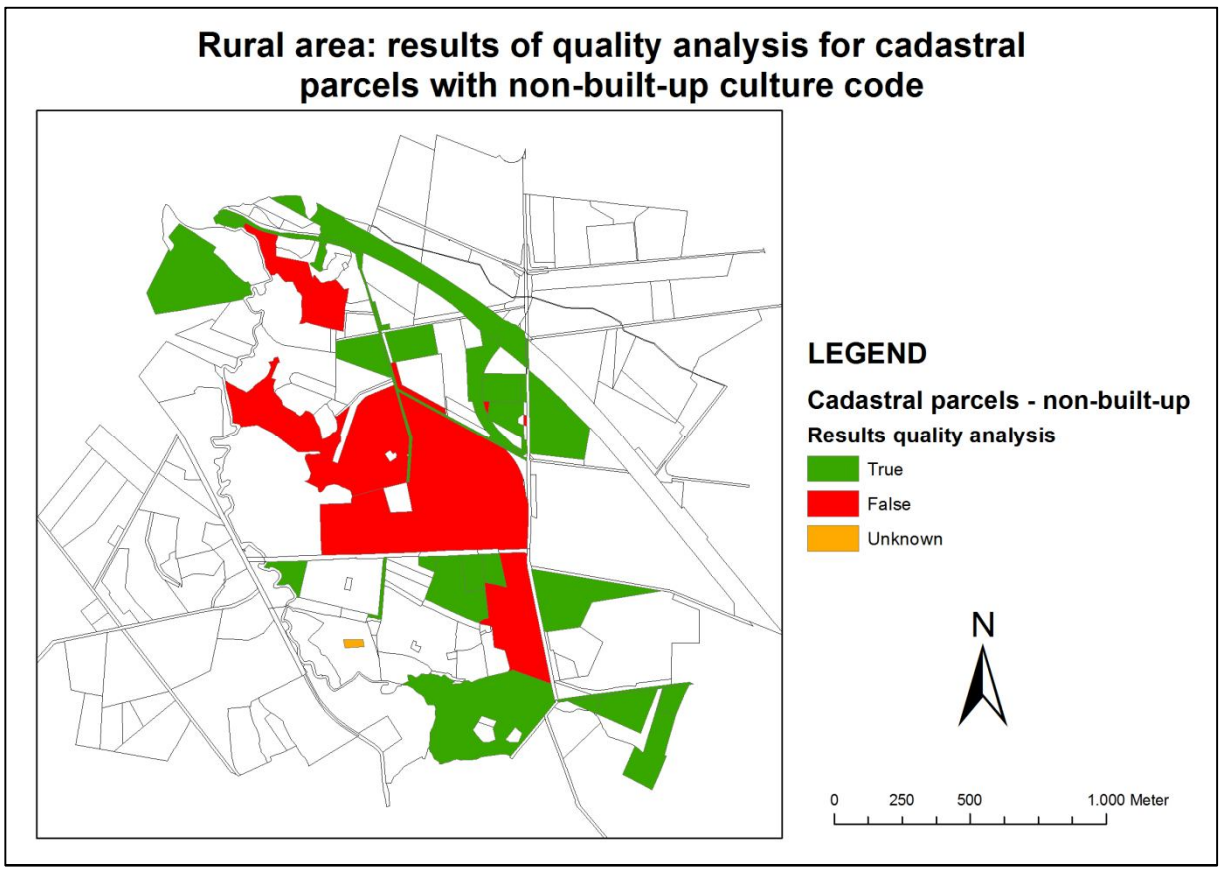
Culture code	Definition
00 Initial	[This culture code is not used in this thesis. Every cadastral parcel has been given one of the other culture codes. If it was not clear which culture code should be given to a parcel, then the parcel has been given the culture code "Special properties"]
11 Residential	"A residential road, area, etc. has only private houses, not offices and factories". For the culture codes this means that the cadastral parcel only contains private houses.
12 Residential (apartment)	Apartment = "a set of rooms for living in, especially on one floor of a building". This culture code is given to cadastral parcels containing private houses that consist of apartments.
14 Residential (agricultural)	Agricultural = "used for farming or relating to farming". This culture code is given to parcels that contain both private houses and farms.
18 Storeroom – shelter (garage-shed)	Storeroom = "A room for keeping things in while they are not being used". Shelter = "A building designed to give protection from bad weather, danger or attack". Garage = "A building where a car is kept, which is built next to or as part of a house". Shed = "A small building, usually made of wood, used for storing things". This culture code is given to cadastral parcels containing buildings that are used for storing things or buildings that are designed to give protection from bad weather, danger or attack.
21 Activity (office)	Activity = "when a lot of things are happening or people are moving around". Office = "A room or part of a building in which people work, especially sitting at tables with computers, telephones, etc., usually as a part of a business or other organization". This culture code is given to parcels containing offices.
22 Activity (industry)	Industry = "the companies and activities involved in the process of producing goods for sale, especially in a factory or special area". This culture code is given to parcels containing factories/plants or any other industry-related building.
23 Activity (utility)	Utility = "A service which is used by the public, such as an electricity or gas supply or a train service". This culture code is given to parcels containing energy plants, nuclear plants, or other electricity- or gas-related buildings. Train services are part of the "Public transport" culture code.
25 Land (industry)	Land = "An area of ground, especially when used for a particular purpose such as farming or building". This culture code is given to land used for industrial activities.
27 Activity (retail trade)	Retail = "The activity of selling goods to the public in stores, on the internet, etc., rather than selling to stores, other businesses, etc.".

	This culture code is given to parcels containing shops, or other retail-related buildings.
28 Activity (catering)	Catering = "To provide, and sometimes serve, food". This culture code is given to parcels containing bars, restaurants, or other catering-related buildings.
29 Activity (agricultural)	This culture code is given to parcels containing agricultural buildings.
34 Parking	"Leaving a vehicle in a particular place for a period of time". This culture code is given to parcels containing parking garages, parking lots, or other parking-related objects.
35 Defense	"Protection or support against attack, criticism or infection". This culture code is given to parcels containing military buildings, military areas, or other defense-related objects.
37 Residential with activity	This culture code is given to cadastral parcels containing a combination of private houses and buildings used for activity (retail trade, office, industry, etc.).
41 Roads	Road = "A long hard surface built for vehicles to travel along". This culture code is given to parcels containing roads, but not roads only used by pedestrians and cyclists or for horse riding.
44 Public transport	"A system of vehicles such as buses and trains which operate at regular times on fixed routes and are used by the public." This culture code is given to parcels containing public transport objects such as railroads, stations and bus lanes.
46 Air traffic	This culture code is given to parcels containing objects such as starting lanes (for airplanes) and airports.
47 Pipes – Tubes	Pipe = "A tube inside which liquid or gas flows from one place to another". Tube = "A long hollow cylinder made from plastic, metal, rubber or glass, especially used for moving or containing liquids or gases". [This culture code is not used in this thesis, as the source datasets lack information on pipes and tubes]
51 Recreation – Sport	Recreation = "(a way of) enjoying yourself when you are not working". Sport = "A game, competition or activity needing physical effort and skill that is played or done according to rules, for enjoyment and/or as a job". This culture code is given to parcels containing recreational or sport objects, such as bungalow parks, stadions, and tennis parks.
53 Residential (recreation)	This culture code is given to parcels containing private houses that have a recreational function (e.g. summer-houses).
55 Parks – Public gardens	Park = "A large area of land with grass and trees surrounded by fences or walls, which is specially arranged so that people can walk in it for pleasure or children can play in it". A public garden is like a park, but normally a bit smaller. This culture is given to parcels containing parks or public gardens.
57 Yard – Garden	Garden = "A piece of land next to and belonging to a house, where flowers and other plants are grown, and often containing an area of grass". A yard is closely related to a garden. This culture code is given to cadastral parcels containing private houses and land next to it.
61 Land (nature)	Nature = "All the animals, plants, rocks, etc. in the world and all the features, forces and processes that happen or exist independently of people, such as the weather, sea, mountains, reproduction and growth". This culture code is given to cadastral parcels containing nature-objects such as forests, sand areas, and dunes.

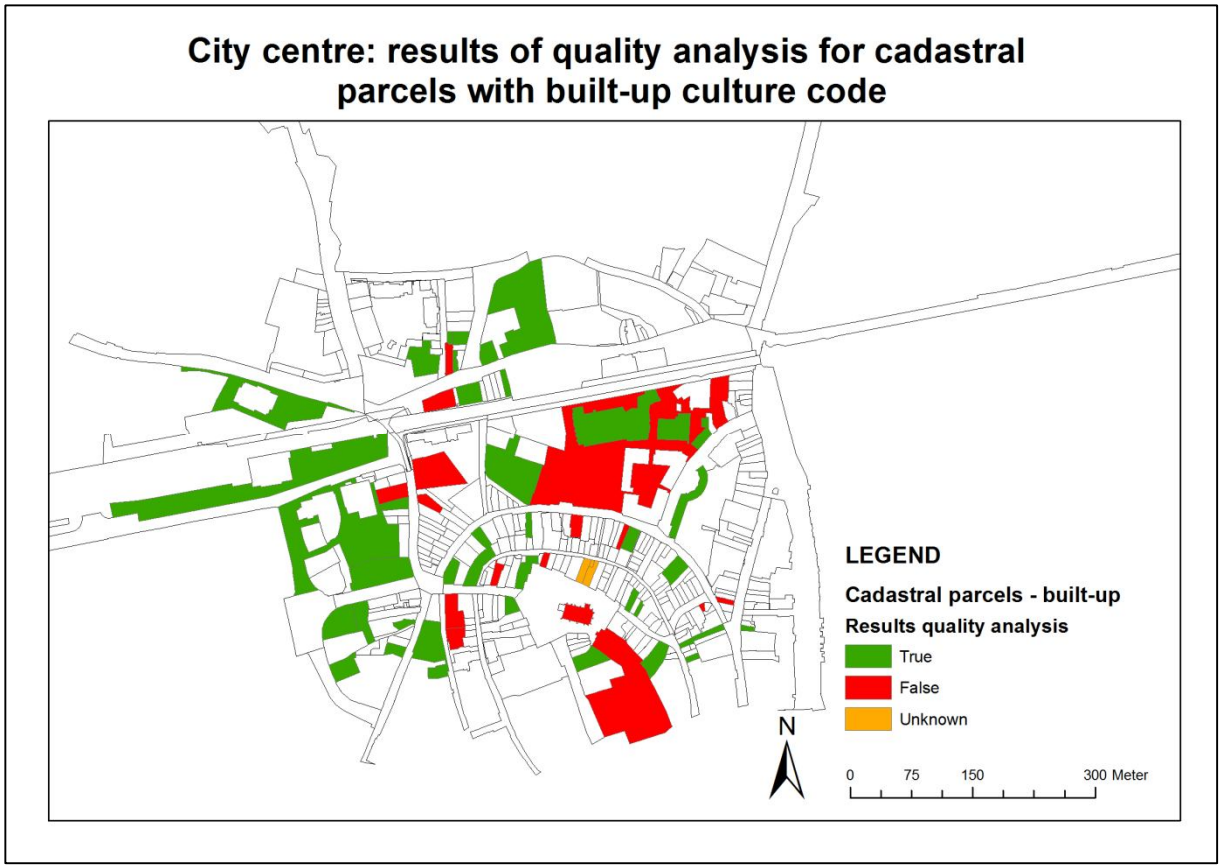
62 Land (farming)	Farming = "The activity of working on a farm or organizing the work there". This culture code is given to parcels containing farming land such as cropland.
63 Land (grassland)	Grassland = "A large area of land covered with grass".
64 Activity (glasshouse)	Glasshouse = "A large building with glass sides and roof for growing plants in". This culture code is given to parcels containing glasshouses.
66 Land (cultivation)	Cultivation = "To prepare land and grow crops on it, or to grow a particular crop". This culture code is given to parcels containing objects such as tree nurseries and bulb-growing land.
72 Health	This culture code is given to parcels containing objects related to healthcare (= "the set of services provided by a country or an organization for the treatment of the physically and the mentally ill"). This culture code is given to parcels containing objects such as hospitals and clinics.
74 Education	"The process of teaching or learning in a school or college, or the knowledge that you get from this". This culture code is given to parcels containing buildings such as schools and universities.
76 Culture	"The way of life, especially the general customs and beliefs, of a particular group of people at a particular time". This culture code is given to parcels containing cultural objects, such as monuments and (open air) theatres.
77 Religion	"The belief in and worship of a god or gods, or any such system of belief and worship". This culture code is given to parcels containing objects such as churches, chapels and monasteries.
78 Police – Fire service	Police = "The official organization that is responsible for protecting people and property, making people obey the law, finding about and solving crime, and catching people who have committed a crime". Fire service (fire brigade = "An organization that is in charge of preventing and stopping unwanted fires"). This culture code is given to parcels containing a police of fire service office and/or garage.
79 Justice	"Fairness in the way people are dealt with". In relation to the culture codes, justice refers to objects such as courts and prisons.
85 Port	"A town by the sea or by a river which has a harbor, or the harbor itself". This culture code is given to parcels containing a harbor and/or objects such as landing-stages/piers and docks.
89 Water	This culture code is given to parcels containing all kinds of water objects, such as rivers, sea and lakes.
92 Land (new construction activity)	Construction = "The work of building or making something, especially buildings, bridges, etc.". This culture code is given to parcels on which a building is built which will be used for activity purposes.
93 Land (new construction residential)	This culture code is given to parcels on which a building is built which will be used as private house(s).
94 Funeral service	Funeral = "A (usually religious) ceremony for burying or burning the body of a dead person". This culture code is given to parcels containing objects such as crematories and cemeteries.
99 Special properties	This culture code is given to those parcels that cannot be given one of the abovementioned culture codes and to parcels that have a combination of non-residential buildings (e.g. office and accommodation).

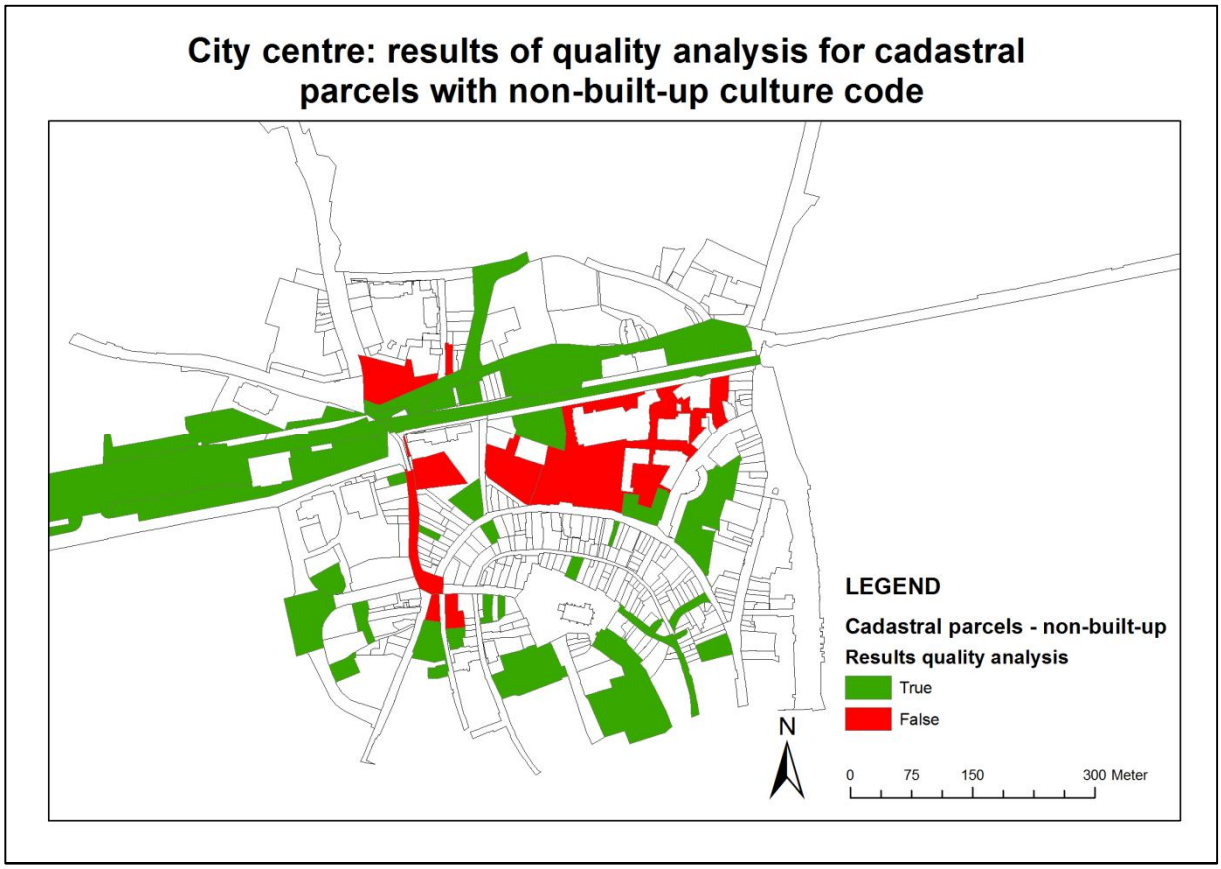
Appendix D – Results quality analysis rural test area



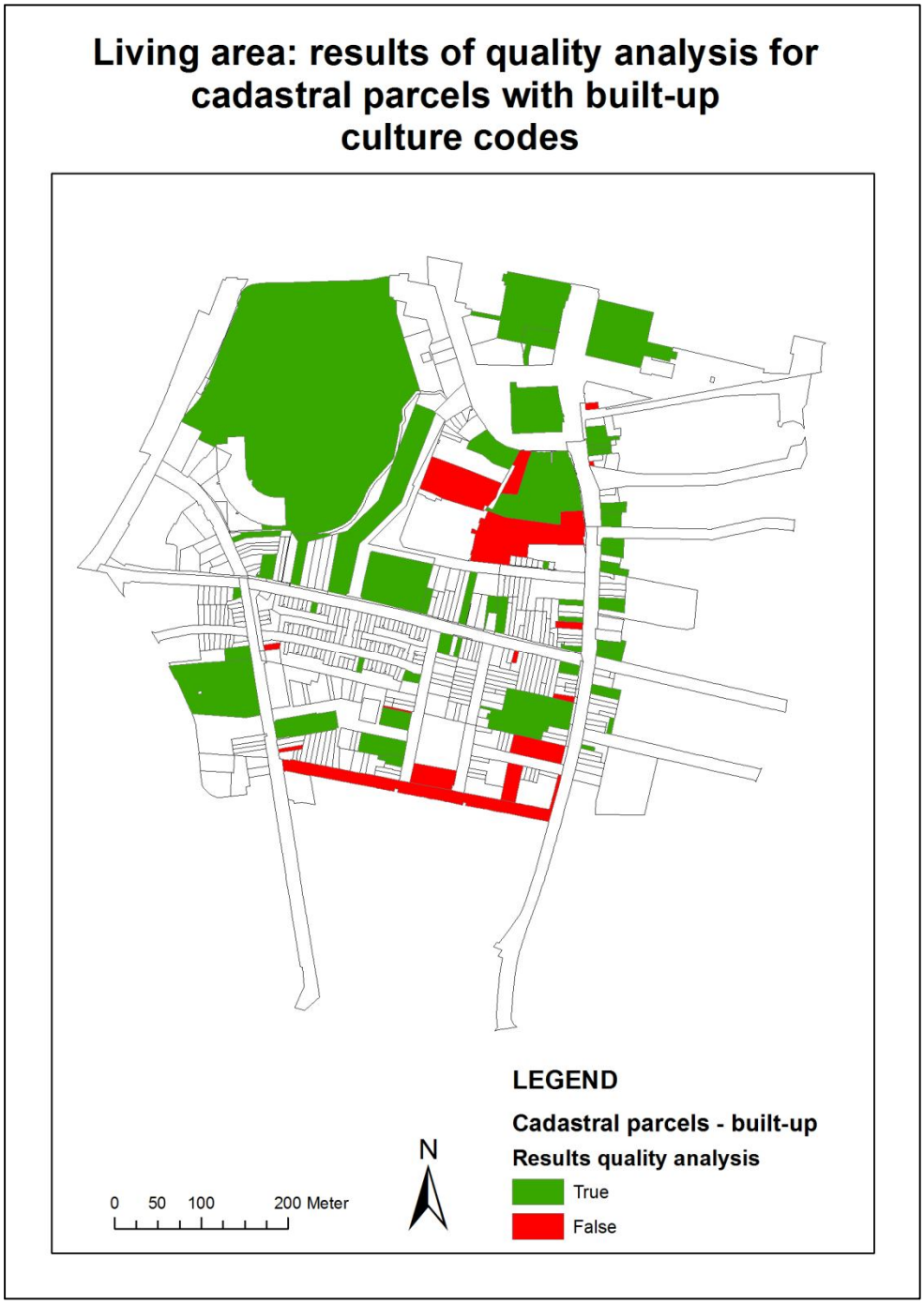


Appendix E – Results quality analysis city centre test area

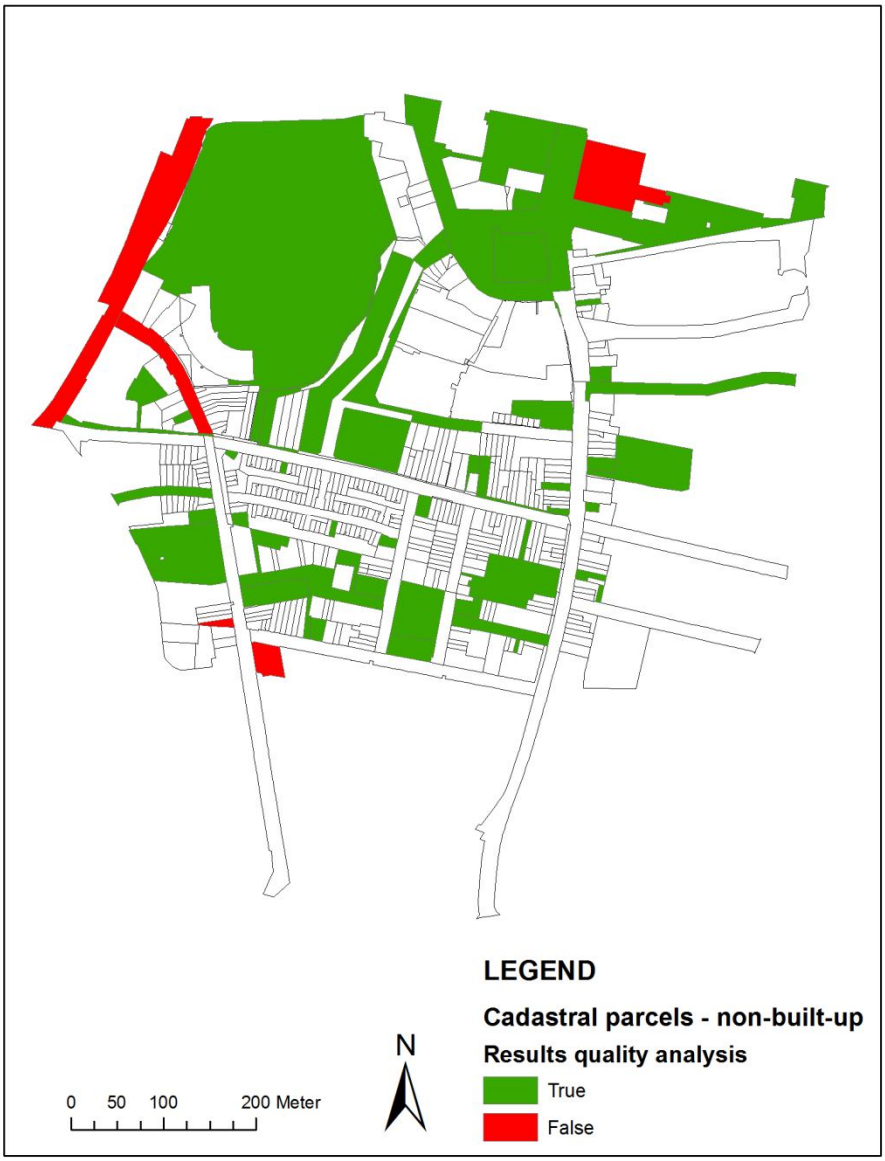




Appendix F – Results quality analysis living area test area



Living area: results of quality analysis for cadastral parcels with non-built-up culture codes



Appendix G – Results quality analysis industrial/harbor test area