The Effect Of Spatial Generalisation On Filtering Noise For Spatio-Temporal Analyses

Anna van Paddenburg and Monica Wachowicz

Wageningen UR, Centre for Geo-Information, The Netherlands
annavp@compuserve.com
M.Wachowicz@Alterra.wag-ur.nl

Abstract. Spatial data sets do not only contain true information, there is also a certain amount of 'noise' associated with the data. The use of these data in spatio-temporal analyses, often results in a sub-optimal representation of reality. Generalising spatial data sets collected at different times may serve the purpose of filtering noise so that spatio-temporal change can be better elucidated. In this paper we aim to test that proposition by addressing the following questions. Does generalisation have a significant influence in the state of the noise in space-time dimensions? Can noise be filtered by a generalisation process? Does it result in a greater probability of detecting environmental variation over time? In the first part of the paper, the main aspects of the generalisation process are presented. The field representation (raster model) is described by three elements of analysis; resolution, spacing and extent. Based on these elements, five generalisation methods are analysed. Following an understanding of these methods, the generalisation process is implemented using different land-use data sets obtained from the classification of satellite images, which are then compared at two different times for spatio-temporal analysis. The observed spatio-temporal variations are presented for each method and the filtering of noise is discussed. The importance of deciding which generalisation method to use for spatio-temporal analysis is highlighted. Results show that noise filtration does occur in the generalising process. This may prove that generalising data for spatio-temporal analyses is beneficial to the quality of the results. As noise is filtered, the observed spatio-temporal variation, after the generalisation process, is probably more representative of the true spatio-temporal change in the real world.

1. INTRODUCTION

The use of Geographic Information Systems and Earth Observation Systems has lead to important advances in research and operational applications in environmental sciences. Merging these two technologies has resulted in a tremendous increase in the generation of environmental data for many kinds of users, such as scientists, policy makers, environmental management bodies etc.. The combined use of these technologies has now reversed the situation where data were sparse to a situation where data are abundant. However, with this abundance of data there is the inevitability of having noise associated with it. In most environmental data sets, noise is generated by errors. Some of these errors represent random fluctuations caused by the measurement process of remote sensing techniques, which may for example be due to atmospheric effects or localised cloud. Other errors may occur when the environmental variable of interest (e.g. categorical variables such as land use classes) is estimated from the measure of remotely sensed radiation. All these errors, and many more, can create noise associated to environmental data, which is then carried through the process when obtaining geo-information from this data set.

Previous studies have shown that spatio-temporal analyses of spatial-data contain a lot of noise resulting in a sub-optimal representation of reality (Sonneveld, *et al*., 2000; Zeeuw and Hazeu, 2001; Wachowicz, 2000; Langran, 1992). Noise is often (un-)intentionally left aside, however critical it is to the final results of spatio-temporal analyses. As technology provides more and more data we need to be increasingly concerned about its quality. Little insight and concern exists about the effects of noise on the quality of the obtained results. Therefore, it is important to eliminate or at least minimise noise in environmental data sets, particularly because both space and time dimensions are involved in analysing spatio-temporal variation; the difference in geo-information between spatial data sets of different times. It is a static representation of change in the real world (usually at two times). The following equations clarify the importance of minimising noise in detecting spatio-temporal variation. If t = time, o = observed measurement, m = true measurement and n = noise, then

at t_1; $\quad o_1 = m_1 + n_1$

and at t_2; $\quad o_2 = m_2 + n_2$

The spatio-temporal variation between t_1 and t_2 is not the difference between the observed measurement (o_1-o_2), but the true measurement (m_1-m_2). Therefore, n_1 and n_2 need to be eliminated or at least minimised to get a good estimate of the true spatio-temporal variation measured by m_1-m_2. For example, when monitoring land-use the true variation over time is the quantity of importance, not the observed variation obtained from a GIS analysis. However, as errors will always be present because of the way that land-use data is measured, collected and analysed it is simply not possible to only obtain true variation. The spatio-temporal variation one gets from an analysis will always contain a certain amount of noise. It is important that this noise is minimised so that the observed variation (o_1-o_2) will be more representative of the true variation (m_1-m_2).

Until now, fundamental research has been focused mainly on predictive spatial data analysis (e.g. Schoorl *et al*., 1997, Fabbri and Chung, 1996, Roughani and Wackernagel, 1990), rather than concentrating on filtering noise in existing geo-databases so that the spatio-temporal change can be better elucidated. As noise is associated with environmental data at cell-level, a change in cell size may affect the associated noise. By generalising spatial cells, the observed spatio-temporal variations may be affected. If this increase in cell size results in a filtering of noise, it will indeed be useful to the observed spatio-temporal variation (o_1-o_2) as its value will be closer to the true spatio-temporal variation (m_1-m_2). For example, if noise is filtered in a spatio-temporal analysis for land-use monitoring, the observed spatio-temporal variation of the land-use will be more representative of the land-use change in the real world. The reason for this generalisation is purely to detect noise in the observed spatio-temporal variation, while maintaining the geo-information. Therefore, it is important that during the generalisation process no significant loss of geo-information should occur. The choice of the size of the cell is extremely important because variations that occur within the space-time dimensions of the cell will not be registered by either the data or the process (Burrough and McDonnell, 1998).

Generalisation in GIS results in a less precise representation of the original data source, however this does not necessarily imply a degraded representation of reality. It enhances certain aspects and reduces others for the purpose of its application. Generalisation can be understood as a process which realises transitions between different models representing a portion of the real world at decreasing detail, while maximising geo-information with respect to a given application (Weibel and Dutton, 1999).

Unfortunately at present, satisfactory implementations of all the transformation operations necessary to achieve comprehensive automated generalisation largely remain to be developed (Weibel and Dutton, 1999). Few have a clear prospective of what the objectives of generalisation should be (Müller *et al*., 1995). Some researchers developed methodologies for vector-based GIS (e.g van Oosterom, 1995, Bundy *et al*.,1995, Buttenfield, 1995, João, 1995), however little research work has focused on field representation-based GIS, especially within a space-time dimension. This paper invites the reader to see spatial generalisation in the light of the resolution of the cell, the spacing between the cells and extent of the area of study.

In the following section (section 2.1), these elements behind the generalisation process are presented. Following an understanding of the decision rules for generalisation (section 2.2), the methods for generalisation (section 2.3), and the spatio-temporal concept (section 3), the generalisation process is implemented using different land-use data sets obtained from the classification of satellite images described in section 4.1. These are compared at two different times for spatio-temporal analysis. The observed spatio-temporal variations are presented for each method. The results and filtering of noise are discussed in section 5. An overall conclusion is presented in section 6 and possible further studies are recommended in section 7.

2. THE GENERALISATION PROCESS

McMaster and Shea (1992) have defined a conceptual framework of digital generalisation, which highlights three areas of generalisation: namely, why, when and how to generalise. The reason why we generalise land-use data in this research is to get spatio-temporal variations closer to reality. How this process is done will be described below. The process of generalisation requires an understanding of the actual 'process' itself and understanding the decision rules used to merge data values. The fundamental elements of analysis in the generalisation process are presented in section 2.1 followed by the choice of decision rules in section 2.2. Based on these principles, the methods for generalisation are discussed in section 2.3.

2.1 Elements Of Analysis

The generalisation process can be described by three elements of analysis within a raster model: resolution, spacing and extent (Figure 1) (McBratney, 2001). This process involves changing the resolution and/or the spacing of spatial data to represent a specified area.

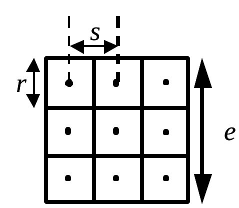


Figure 1: The r, s and e of a raster model

Resolution (*r*) refers to the cell size (this is measured by for example the length on the physical ground). In raster models the cells are always assumed to be square. Spacing (*s*) refers to the distance on the physical ground laterally or horizontally between the centres of cells. In raster models, the case is often that *r* = *s*. Extent (*e*) refers to the size of the study area represented by the raster model. The generalisation process can be achieved by:

(i) manipulating *s* and holding *r* constant (Figure 2.1). By increasing *s* not all values of the cells are used in the process. In Figure 2, *s* is trebled, this results in nine possible images that can be taken further into the process but only 1/9th of the information in the block is used. This process can be considered the same as sub-sampling data values, taking only 1/9th of the values further in the process and giving all cells within the block that same value.

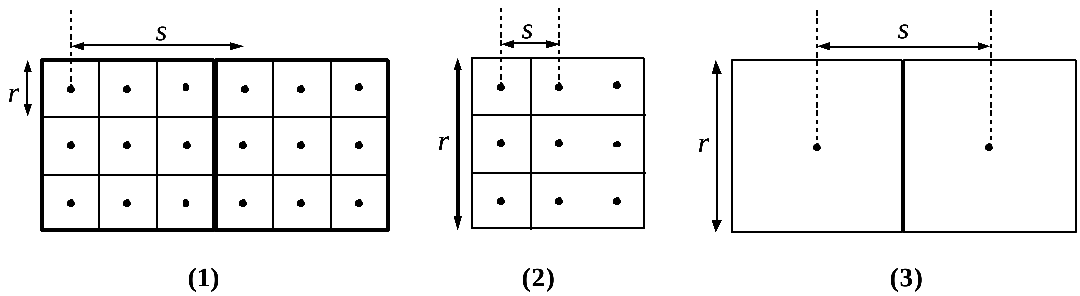


Figure 2: Three ways to generalise raster models

(ii) manipulating *r* and holding *s* constant (Figure 2.2). This involves increasing the area that the cell represents. When *r* is increased the resolution of the cell is decreased but the spacing between the cells remains the same.

(iii) manipulating both r and s, often done to maintain *r* = *s* (Figure 2.3). This is similar to (i) since only 1/9th of the information in the block is taken. However, the difference is that *r* is also increased, so what is represented by 9 cells in (i) is here represented by 1 cell. The size of the generalised area is formed based on the magnitude of *r* or *s*. This size determines the block size of the generalised area. A cell that is generalised in the way that *r* is manipulated and *s* is held constant, is referred to as a focal method (Figure 3.1), In focal methods each cell undergoes a generalisation process resulting in overlapping blocks. The block size is defined by the resolution of the cell. The process whereby a cell is generalised by manipulating *s* and holding *r* constant is referred to as a block method (Figure 3.2). Here, only a subset of the data is looked at and the block size is defined by the spacing between the cells. The cells within the block will all get the same value.

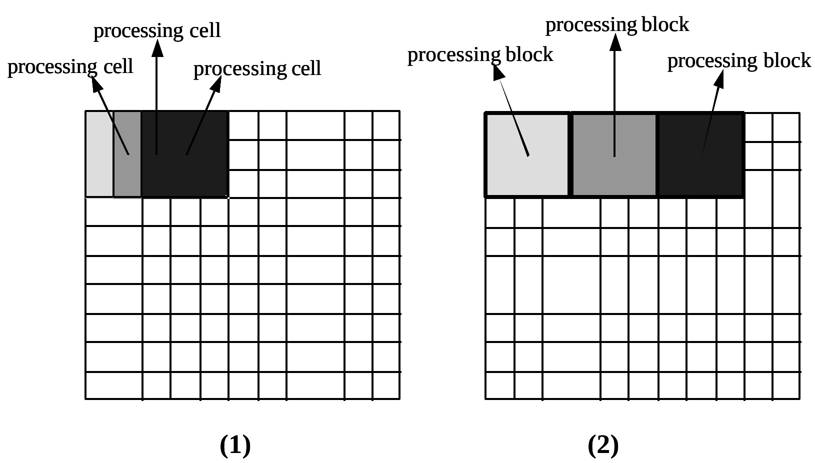


Figure 3: An illustration of the Focal and Block Generalisation Methods

2.2 Decision Rules For Generalisation

There are many decision rules to merge data values in cells. For the purpose of detecting low-level widespread spatio-temporal change, such as land-use changes, three types of decision rules are proposed to be incorporated into the generalisation process. When cells describe discrete data values (e.g. land-use classes) the **majority** rule is normally used to derive the new data value for the generalised cell (Star and Estes, 1990). A rule commonly used to generalise continuous data values (e.g. the probability or the proportion of grass) is by calculating the **mean** of the data values in the specified block (Star and Estes, 1990). This decision rule can also be implemented on discrete data such as land-use, however only one land-use type can be analysed at a time. In this case the particular land-use type (e.g. grass) should be given the value 1 and remaining cells (e.g. not grass) the value 0. The last decision rule to be incorporated in this paper is by calculating the **weighted average** of the data values. Similar to the mean decision rule, only one type of land-use can be analysed at the time when this rule is applied to land-use data. The weighted average is calculated using Block Kriging, whereby the estimated average is based on the strength of the correlation between the data values of the cells, as a function of the distance between the cells (Burgess and Webster, 1980).

2.3 Methods For Generalisation

Five methods can be applied according to the elements of analysis and the decision rules of a generalisation process. The proposed generalisation methods can manipulate *r* and hold *s* constant (focal method), or hold *r* constant and manipulate *s* (block method). In its process it can use the majority, the mean or the weighted average decision rule to merge data values.

Based on these principles, five methods for generalisation are proposed;

1. Majority focal method (*r* is manipulated, *s* is constant)
2. Majority block method (*r* is constant, *s* is manipulated)
3. Mean focal method (*r* is manipulated, *s* is constant)
4. Mean block method (*r* is constant, *s* is manipulated)
5. Weighted average focal method (*r* is manipulated, *s* is constant)

2.3.1 Majority Focal Method

This method is processed on a cell-by-cell basis. In this process, *r* is increased (resolution is decreased), whereas *s* is constant. The process gives a new value to each cell based on the values of the surrounding neighbourhood which size is defined by the size of *r*. The majority decision rule is applied to discrete data values, such as land-use classes. If grass is considered the land-use class of concern, then the cell could either be the class ‘grass’ or ‘not grass’. During the process, the class that covers the largest fraction within the block size determines the class of the new generalised cell. If *r* is increased 3 times, the generalised area covers a block of 3 by 3 cells. The new value of the processing cell is the class that covers the largest fraction in the 9 surrounding cells. The example in Figure 4.1 shows that the original data set contains 3 cells with the class ‘grass’ and 6 cells with the class ‘not grass’, therefore after the generalisation process the processing cell in the centre of the block is given the class ‘not grass’. As the spacing is kept constant, data values in each surrounding cell is used when determining the new value of each processing cell.

2.3.2 Majority Block Method

Block methods are processed per block, where only a subset of the data is of interest. During this process, *s* is increased whereas *r* is held constant. The process gives a new value to each block based on the values of the surrounding neighbourhood which size is defined by the magnitude of *s*. The surrounding neighbourhood consists of cells that fall within the block. This method also follows the majority decision rule. The value that covers the largest fraction of the block determines the value of the generalised area, as described in the previous method. Figure 4.2 illustrates that after the generalisation process all cells in the block get the same value.

2.3.3 Mean Focal Method

The generalisation is processed on a cell-by-cell basis (*r* is increased whereas *s* is held constant). The mean decision rule finds the average of the block. This cannot be done directly on thematic spatial data. However, when ‘grass’ is given the value 1 and ‘not grass’ the value 0, it is possible to find the fraction of grass within the block (Figure 4.3). The information in the processing cell indicates that a third of the proportion of the information contained in the generalised cell is ‘grass’.

2.3.4 Mean Block Method

This generalisation process is processed per block, similar to the majority block function only a subset of the data is of concern. The difference lies in the decision rule. The process is carried out based on the mean of the values in the cells of the block. The cells within this block will get the same new value, which is the average of the total values of the cells. Figure 4.4 illustrates the generalisation process of the land-use grass. After generalisation the information in the cells indicate that a third of the information contained in this block is grass.

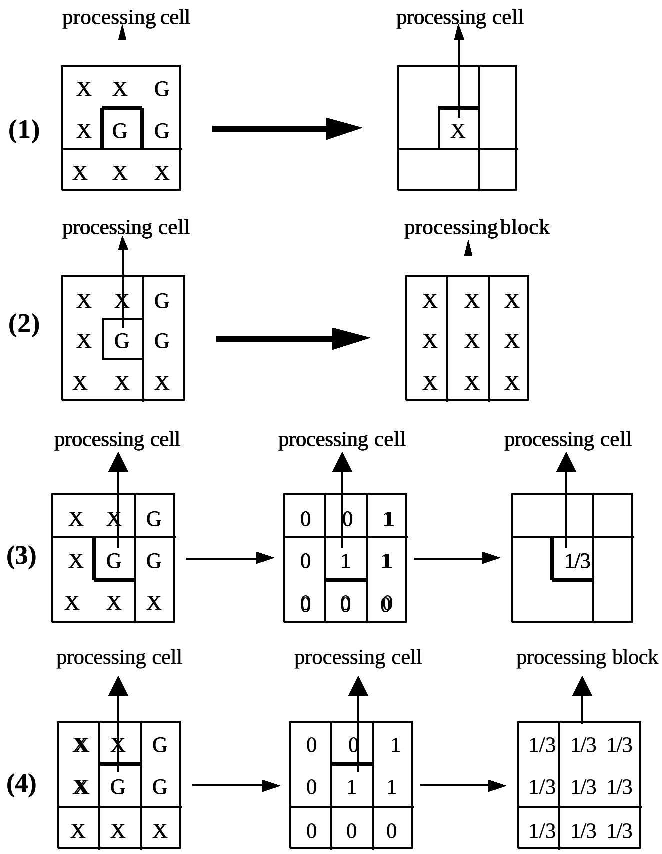


Figure 4: The majority and mean decision rules applied on focal and block methods

2.3.5 Weighted Average Focal Method

This process follows the focal method where every cell is processed. The new generalised value is obtained by calculating the weighted average of each cell. Block Kriging is used to find the weighted average and unlike other methods uses data from outside the block to get the new value for a cell. In all other methods, the surrounding neighbourhood of the processing cell is as large as the block. In Block Kriging the surrounding neighbourhood is larger than the block.

The principle behind Kriging is that the estimated new value is a linear weighted average of proximate sample data. The method is based on estimating the strength of the correlation between the data points, as a function of the distance between the points (Burgess and Webster, 1980). Kriging is a linear technique for optimal unbiased estimation (Curran *et al*., 1998). It is unbiased because the weights are forced to sum to one and it is optimal in the sense that the estimation variance s_E^2 or its square root, the standard error, is minimised (Curran *et al*., 1998).

The weighted average is obtained by estimating the average based on the distance between the points, within a neighbourhood of around 5 cells either side of the processing cell. This will provide 100 points for the estimation of the variogram and for computing the average. The weights, which are applied to each cell to form the new value, depend on the variogram and the configuration of the data. If the variogram function increases with distance, cells closer together are similar to each other and cells further apart are more dissimilar, therefore the weights are larger for cells nearest to the processing cell. The weights of the block sum to 1. The new value is formed by multiplying the weight at each of the cells in the block by the data values in each of the cells in the block and adding them up. Because the weights sum to one and the data values are either 0 or 1, the resulting number is a value between 0 and 1, which is the weighted average of the processing cell. The output of Block Kriging provides information about the weighted average of each cell and the uncertainty in each cell.

Figure 5 shows that the weights outside the block take part in the final decision process of the generalisation. Depending on the variogram, the values in each cell will be assigned a certain weight which will influence the final weighted average of the processing cell. As the example in figure 5 illustrates the presence of more cells with the class ‘grass’ than ‘not grass’ in its surrounding neighbourhood, the weighted average of the processing cell will result in a higher value (which could be more representative when the value of the cells are used for spatio-temporal analysis).

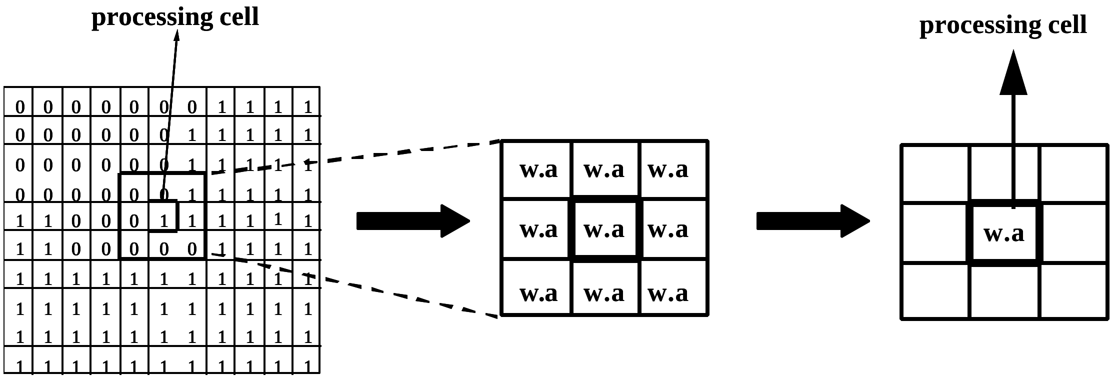


Figure 5: An illustration how the weighted average rule is applied to the focal method

3. SPATIO-TEMPORAL ANALYSIS

Despite having interrelated aims, research in temporal and spatial database models has predominantly developed independently (Wachowicz, 2000). As a result, no temporal geographic information systems are yet readily available for spatio-temporal analysis such as monitoring land-use changes. Although maps may be used to depict change, they have mainly been used to present a static view of the world. Since changes and time are dependent, time and change are strongly related. It is possible to store the changes in a database as an approach to reason about time. Therefore, there are two approaches to do temporal analyses: a change-based approach and a time-based approach (Al-Taha and Barrera, 1990). The change-based approach concentrates on recording changes or facts valid at a certain point in time and the time-based approach considers time as a separate dimension, similar to a one-dimensional space.

Wachowicz (1999) clarifies the fundamental differences between the two approaches. The concept behind the change-based approach is the *absolute view of space*, which considers space as finite, homogeneous, and isotropic, with an existence fully independent of any entity it might contain. In space-dominant data models, time is implicitly incorporated into the spatial data model every time a change occurs. When time takes part explicitly in a model (the time-based approach), either with or without reference to space, the time dominance is generated and an *absolute view of time* is used within a model. In time-dominant data models, time is represented by a fourth dimension, a time line marked out with intervals, along which events can be located.

Although time is generally perceived as continuous, the preference for a discrete time representation stands out in space-dominant models. Time is treated as a discrete subset of the real numbers ordered linearly. Therefore, changes are supposed to take place a finite number of times so that each change produces a historical state indexed by time. This preference has shaped the way current spatio-temporal analyses are carried out. Most spatio-temporal analyses are based on the change-based approach, where the time dimension is simply added to the spatial dimension. Change is recorded as a series of 'snapshot' images. These 'snapshot' images entail use of a raster data model with a sequence of spatially-registered cells. Each sequential cell represents a 'true state' at a particular moment. The 'true state' of any stored time can be retrieved directly. Nevertheless, the changes that occur between 'true states' are not explicitly stored. These must be calculated by comparing the spatial patterns of two successive states (Peuquet, 1994). It is important to note that adding the time dimension to a spatial data model is inadequate for representing space and time in databases as it will result in a database model that represents the time dimension in the same manner as the spatial dimension. Peuquet (1994) points out that absolute space is objective since it gives us an immutable structure that is rigid, purely geometric and serves as the framework in which objects may or may not change. This might be reason why most of the GIS products adopted the space-dominant view within their database models.

The database used for this study only supports the absolute view of space. As a result the change-based approach is adopted for the spatio-temporal analysis. Two types of analyses are used to obtain the total observed spatio-temporal variation (o_1-o_2) which represents the total change in the real world. The first analysis merely compares the statistics of the spatial images between the different times (regional level analysis) and the second follows the method where change is recorded as 'snapshots' images (cell-level analysis).

(i) Regional level analysis

Using this comparison no information is given to the position of the spatio-temporal variation. Only the statistical appearance at regional level is considered (Zeeuw and Hazeu, 2001). The observed spatio-temporal variation in the total area is independent of location. It is simply the statistical difference between the total areas of the two different points in time.

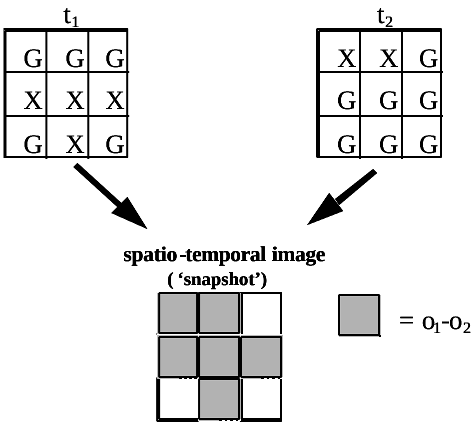


Figure 6: Spatial illustration where the total variation between the regional level analysis is different to the cell-level analysis (G='grass' and X = not grass')

(ii) Cell-level analysis

This is applied to analyses whereby variation at every cell is taken into account (Zeeuw and Hazeu, 2001). The analysis is location dependent. An overlay of the two digital maps will create a 'snapshot' which represents the 'world state' of the stored time. It provides the necessary information about the variation that occurred over that particular time at cell-level. This results in a new map which displays the observed spatio-temporal variations between the two maps. The total area of these variations can then be calculated.

The above example illustrates the differences between the two types of analyses (Figure 6). The statistics at regional level show that the total cells with the value 'grass' is 5 at t_1 and 7 at t_2. Hence, the total observed spatio-temporal variation (o_1-o_2) at regional level amounts to 2 variations. At cell-level however, the creation of a 'snapshot' allows us to locate where these variations occur; changes whereby the cell value has become 'grass' and 'not grass' are both recorded. This results in a total of 6 variations over the specified time.

4. IMPLEMENTATION

4.1 Overview of Approach

The generalisation process is implemented on land-use data (grass). This is to test whether generalisation serves the purpose of filtering noise so that spatio-temporal change can be better elucidated. Two land-use data sets collected at different times are compared before and after the generalisation process. In the generalisation process *r* or *s* is increased three, five and seven times that of the original image. The total observed spatio-temporal variations (o_1-o_2) are compared between the original geo-information and the generalised geo-information. A schematic representation of the research approach is shown in Figure 7.

4.2 Land-use Data

Two sets of data derived from the National Land Use database, LGN (Landelijk Grondgebruik Nederland) data, obtained from satellite images in 1992 (LGN 2) and 1996 (LGN 3), have been used in this study. This National Land-use database has a raster format with cells of 25m x 25m. It was mainly derived from the classification of remotely sensed data sources (e.g. SPOT (20m x 20m resolution) and Landsat (30 m x 30m resolution) and is updated once every four years. The applied classification procedures were partly automatic, partly manual. Classes that were difficult to distinguish from remote sensing data were classified with the aid of GIS techniques from other data sources. For example, LGN 2 classification was refined by combined application of satellite imagery from 1990, 1992 and 1994, information from BARS (Basisbestand Ruitelijke Structuren van de Rijks Planologische Dienst) and with aid of topographical maps, aerial photos, agricultural statistics from the CBS and reference field data (Noordman *et al.*, 1997). The taxonomy used in the classification of LGN 2 and LGN 3 is almost identical, yet LGN 3 consists of 26 classes and LGN 2 consists of 25 classes. To make comparison between the two data sets possible, LGN 3 has been slightly simplified to 25 classes. The class called 'build-up in agricultural areas' has been combined with the class 'grass' as was the case in LGN 2. For the purpose of this research, the study area only covers the area of MAP 32West which covers the municipality of Soest and its environs (50113 ha). In this data set only the class 'grass' is observed. All other classes have been combined and labelled 'not grass'.

Unfortunately, no reference data are available to validate findings in this study. Therefore, an attempt to verify the results an adapted LGN 2 data set has been included in the analysis, which have been obtained by an adapted data capturing technique (Zeeuw and Hazeu, 2001). This technique presents a new data set by deriving geo-information from the spatio-temporal image ('snapshot') of LGN 2 and LGN 3 which shows the observed spatio-temporal variation (o_1-o_2) between the two maps. From this

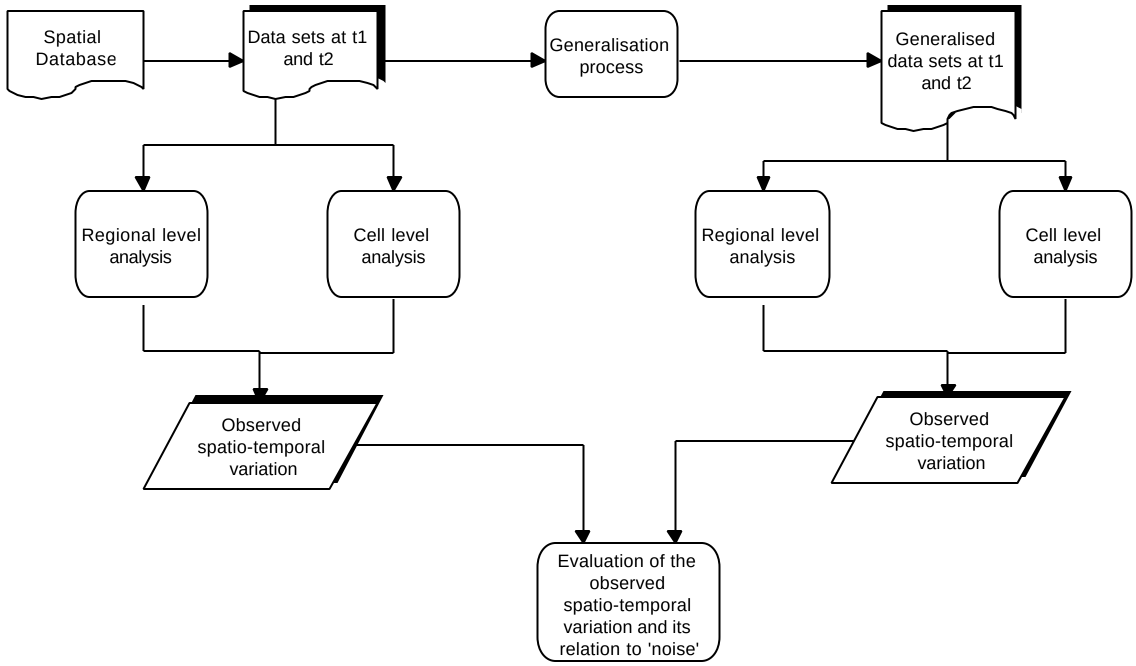


Figure 7: Schematic representation of research approach

spatio-temporal image, the variation between the two data sets can be seen. Cells where change has occurred has then been given values from LGN 3 and unchanged cells have been given values from LGN 2. The observed spatio-temporal variation (o_1-o_2) between LGN 3 and this adapted LGN 2 is used to verify results. The main objective of this data capturing approach is to adapt the data capturing of temporal geographical data in such a way that spatio-temporal analysis of geo-information will result in more reliable results (Zeeuw and Hazeu, 2001). However, Zeeuw and Hazeu (2001) confirmed in their recent study that the spatio-temporal results obtained from LGN 3 and the adapted version of LGN 2 resulted in an underestimation of true spatio-temporal variation (m_1-m_2).

4.3 The Generalisation Process

The five generalisation methods described in section 2 were applied to each data set. For the mean methods and the weighted average method the class ‘grass’ has been given the value 1 and ‘not grass’ the value 0. The new generalised cell value is calculated based on its surrounding neighbourhood (see Figure 4). The result of this process is two new maps at different times with generalised values in the cells. For the block methods where the spacing is increased, all generalised values of the cells within a block are the same. The cells using the majority methods will give the value ‘grass’ or ‘not grass’ in the generalised cells. The cells using the mean methods and the weighted averages will give a value between 0 and 1 (Figure 8).

The majority focal and mean focal methods were applied using the majority rule (under GIS analysis) in Erdas Imagine. The majority block and the mean block methods were applied using the BlockStats request in ArcView (Spatial Analyst). Weighted averages were calculated using a program called VESPER (Variogram Estimation and Spatial Prediction with Error) developed for spatial prediction (Minasny *et al*., 1999). VESPER is an interpolation technique that automatically performs localised geostatistical analysis. For the purpose of this study local ordinary block kriging with a local spherical variogram is used.

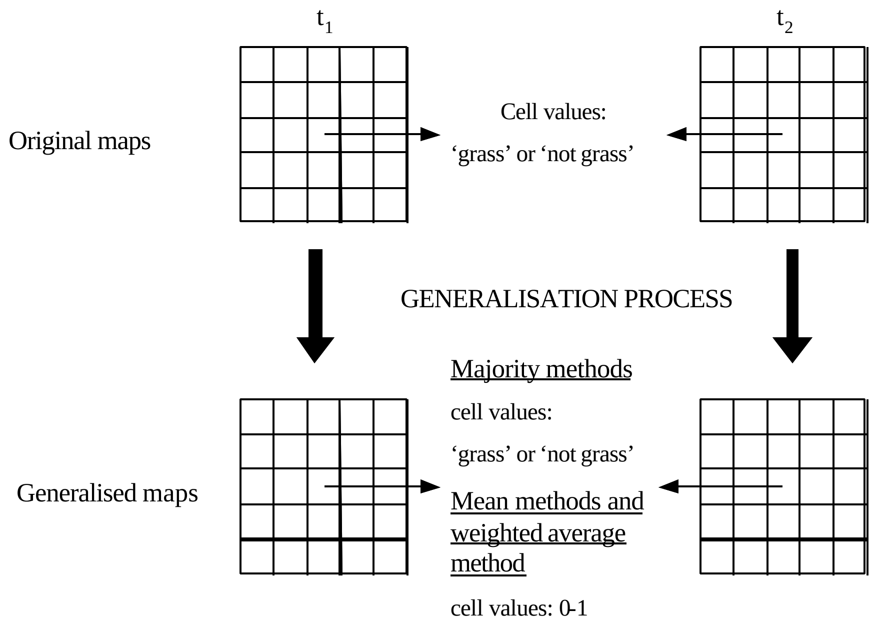


Figure 8: Overview of the generalisation process

4.4 The Spatio-temporal Analysis

(i) Regional level analysis

At regional level, the total area of grass from the majority methods is calculated by adding all cells with the class 'grass'. Cells generalised using the mean methods and the weighted average method however, are made up of values between 0 and 1 because the new value of the cell represents the fraction or the weighted average of the presence of grass in the cell. In this case, the total area of grass is the sum of the fractions or weighted averages multiplied by the size of the cell (25 metres by 25 metres). This is done for both maps at different times and the overall difference between the maps is the total area of the observed spatio-temporal variation (o_1-o_2), which is presented as a percentage of the total area of MAP 32West.

(ii) Cell-level analysis

At cell-level, variation at every cell is taken into account, so with a simple overlay technique a spatio-temporal image ('snapshot') is created, showing the observed spatio-temporal variation between the two maps at different times (Figure 6). Cells with the class 'grass' obtained from maps processed with the majority rule are simply added up. The maps with cell values between 0 and 1 obtained from the mean methods and the weighted average method are subtracted from each other. The resulting image consists of cells with values between –1 and 1 (depending on whether the proportion of grass has increased or decreased over time).

The example above shows that at t_1 0.4 of the area that the cell represents is 'grass' and at t_2 the proportion of 'grass' is 0.7. The value of the spatio-temporal cell is –0.3. This means that over time the proportion of grass has increased by 0.3 (Figure 9).

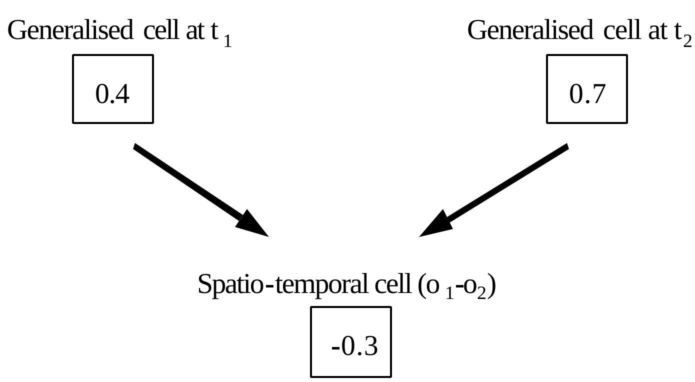


Figure 9: Fictive example of the cell-level analysis for maps generalised with the mean and weighted average methods

The absolute cell values are multiplied by the size of the cell to obtain the total area of the observed spatio-temporal variation. The total area of these variations are presented as a percentage of the total area of MAP 32West.

Similarly, the percentage variation between LGN 3 and the adapted LGN 2 is calculated at regional and cell-level. Since no reference data is available to verify obtained results, this information is used to indicate whether our results are above or below this verifying value which according to Zeeuw and Hazeu (2001) is an underestimation of true spatio-temporal variation (m_1-m_2). If the percentage variation between the two data sets reaches this value, the observed spatio-temporal variation (o_1-o_2) is an underestimation of the true spatio-temporal variation (m_1-m_2).

5. RESULTS AND DISCUSSION

The relation between the observed spatio-temporal variations (o_1-o_2) and the true spatio-temporal variation (m_1-m_2)

All observed spatio-temporal variations calculated using different generalisation methods have values larger than the percentage variation obtained from the verifying data (Table 1 and 2). As the verifying data is an underestimation of true spatio-temporal variation (m_1-m_2), this verifies that the percentage of observed spatio-temporal variation obtained using all generalisation methods and at all proposed magnitudes of generalisation has not passed the underestimated value of 0.88% for the analysis at regional level (Table 1) and 1.95 % for the cell-level analysis (Table 2). Other studies on the LGN database have confirmed that a lot of noise is associated to the data (Sonneveld *et al*.,2000; Zeeuw and Hazeu, 2001). This suggests that the true spatio-temporal variation (m_1-m_2) lies between the variation obtained from the original data sets and the variation obtained from the verifying data sets. So at regional level the true spatio-temporal variation lies between 0.88% and 2.12%, and at cell-level between 1.95 and 5.24%. Due to the absence of a reference data set, no more can be said about the observed spatio-temporal variations regarding its relation to the true spatio-temporal variation (m_1-m_2) in the real world.

Table 1: Regional level analysis showing the percentage of the total observed spatio-temporal variation of ‘grass’ obtained from two data sets at different times using different generalisation methods

Data sets	**% observed spatio-temporal variation (o_1-o_2)**				
	Majority focal	**Majority block**	**Mean focal**	**Mean block**	**Weighted average focal**
Verifying data sets	0.88	0.88	0.88	0.88	0.88
Original data sets	2.12	2.12	2.12	2.12	2.12
Data sets generalised 3 x	2.07	2.10	2.25	2.13	2.24
Data sets generalised 5 x	2.19	2.24	2.16	2.11	2.18
Data sets generalised 7 x	2.20	2.18	2.15	2.10	2.15

Table 2: Cell-level analysis showing the percentage of the total observed spatio-temporal variation of ’grass’ obtained from two data sets at different times using different generalising methods

Data sets	**% observed spatio-temporal variation (o_1-o_2)**				
	Majority focal	**Majority block**	**Mean focal**	**Mean block**	**Weighted average focal**
Verifying data sets	1.95	1.95	1.95	1.95	1.95
Original data sets	5.24	5.24	5.24	5.24	5.24
Data sets generalised 3 x	5.12	4.81	5.03	4.81	5.05
Data sets generalised 5 x	5.08	4.77	4.78	4.57	4.74
Data sets generalised 7 x	4.65	4.20	4.58	4.33	4.50

Observed spatio-temporal variation (o_1-o_2) at regional level

Analysis at regional level only considers the statistical appearance at regional level. No importance is given to the position of the land-use class 'grass'. This kind of geo-information is only important when one is interested in the spatio-temporal variation of the total area. As the differences in the methods are all cell specific, the final spatio-temporal variation of the total area are similar for all methods (Table 1). There is no significant trend apparent in the observed spatio-temporal variation at regional level (Table 1). The total observed spatio-temporal variation of all the generalised data vary between 2.07 and 2.25 %, not much different to the observed spatio-temporal variation of the original data (2.12%). This also indicates that the methods have not influenced the original values.

Observed spatio-temporal variation at cell-level

By creating a spatio-temporal image, it is possible to locate where spatio-temporal variation has taken place, because the variations are analysed at cell-level. The geographical position where changes occurr is recorded (Figure 6). This information is indeed of great importance for monitoring low level widespread change. By creating a spatio-temporal image an indication of what and where changes have happened between two dates may be presented. A decreasing trend is noted in the percentage of observed spatio-temporal variation at cell-level as cells are generalised (Table 2). This decrease accounts for the filtering of noise and the loss of geo-information during the generalisation process.

Assume: P (grass | class = grass) = 0.9 and P (grass | class = not grass) = 0.1

G	X	G
G	G	G
G	G	G

For a block of 9 cells:

E (grass) = P(1-P) = 0.36

Var (fraction) = 9 x $(1/9)^2$ x P (1-P) = 0.01

G	G	X	X	X	X	X	X	X
X	G	G	X	X	X	X	X	X
X	X	G	G	G	X	X	X	X
X	X	G	G	G	G	X	X	X
X	X	X	G	G	G	G	X	X
X	X	X	X	G	G	G	G	X
X	X	X	X	X	G	G	G	G
X	X	X	X	X	X	G	G	G
X	X	X	X	X	X	X	G	G
X	X	X	X	X	X	X	G	G

For a block of 90 cells:

E (grass) = P(1-P) = 0.36

Var (fraction) = 9 x $(1/90)^2$ x P(1-P) = 0.001

Figure 10: The effect of spatial generalisation on noise illustrated by a sequence of independent Bernouli experiments

(i) The effect of spatial generalisation on noise

As we suggested at the start of the paper, data in a cell does not only represent true information but also includes noise associated to the data. When a cell is generalised, its new value represents a greater area with more associated noise. However, if the block size is an area of 9 cells, the new value of the generalised cell will only carry through $1/9^{th}$ of the variance due to noise associated to the area and if the block is an area of 90 cells, the new value of the generalised cell will only carry through $1/90^{th}$ of the variance due to noise associated to it. Hence, a filtering of noise occurs as cells are generalised. This is illustrated by a sequence of independent Bernouli experiments where noise is assumed to be spatially independent (Figure 10) The expectation of a fraction of grass in a block of 9 cells is the same as the expectation of a fraction of grass in a block of 90 cells when the proportion grass : not grass is equal. However, the variance per cell differs. In the block of 9 cells the variance per fraction is 0.01 whereas the variance per cell in the block of 90 cells it is 0.001. The difference in the variance explains the filtering of noise when the cells are blocked. Therefore, relating this back to spatial generalisation, as the *r* or the *s* increases, the averaging covers a larger area allowing more filtering of noise. Generalising 1 cell from 90 cells carries through less noise (var = 0.001) than generalising 1 cell from 9 cells (var = 0.01). The filtering of noise is therefore due to taking the average of more cells. By filtering noise in the generalisation process the observed spatio-temporal variation will be closer to the true spatio-temporal variation. This phenomenon occurs for every generalisation method proposed. Less noise is present in the generalised data sets. This filtering of noise may result in a greater probability of detecting change in spatio-temporal analysis.

(ii) The effect of spatial generalisation on the geo-information content

Spatial generalisation was done by either manipulating the resolution of the cell (focal method) or the spacing between the cells (block method). In our results, a steeper decrease in the percentage of the total observed spatio-temporal variation is observed when generalisation was done using the block method (Figure 11 and 12).

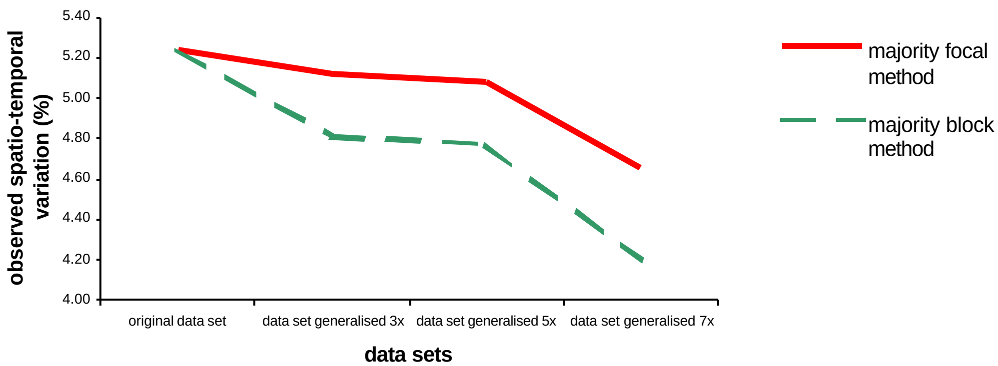


Figure 11: Relation between the observed spatio-temporal variation and generalisation using the majority focal and block methods

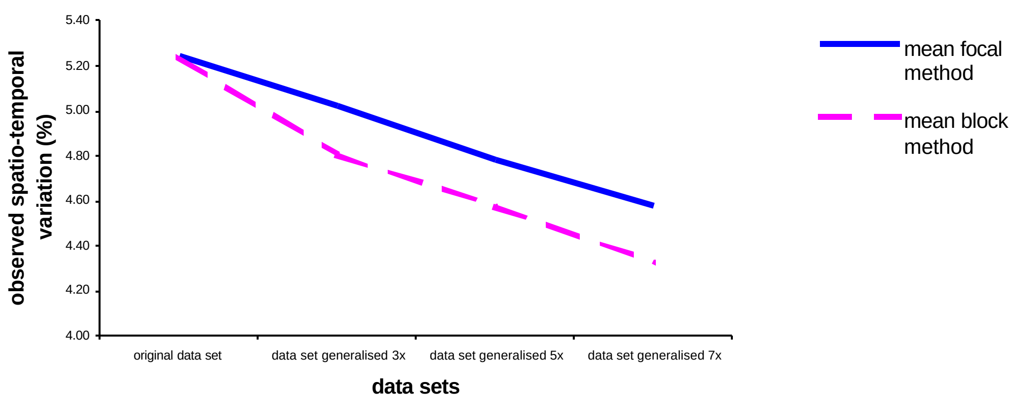
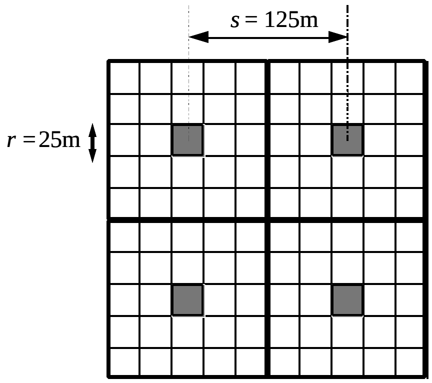


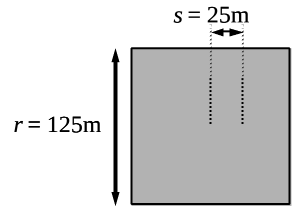
Figure 12: Relation between the observed spatio-temporal variation and generalisation using the mean focal and mean block methods

The percentage of the total observed spatio-temporal variation of the original data sets is 5.24% (Table 2). When generalised 7x the total observed spatio-temporal variation is 4.20% (using the majority block method) and 4.33% (using the mean block method) (Table 2). Figures 11 and 12 clearly indicate that after generalisation the percentage variation obtained by the block methods is lower than the percentage variation obtained by the focal methods. This suggests that the decrease in spatio-temporal variation when manipulating the *s* during the generalisation process is most likely due to a loss of geo-information.

For spatio-temporal analysis it is very important to keep as much information as possible at cell-level so that detailed and representative geo-information is available for monitoring. True variations that are present between the two times and not recorded at cell-level will not be recorded in the spatio-temporal analysis either. Therefore, it is important to maintain as much geo-information from the original data source. The decreasing trend of the observed spatio-temporal variations as generalisation occurs is therefore partly due to a loss of geo-information. Some generalisation methods are better at maintaining geo-information than others.



(1)



(2)

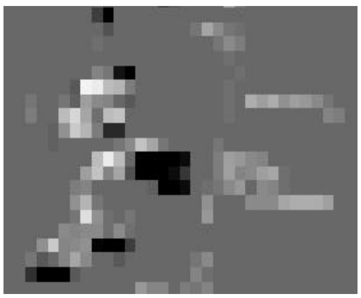
Figure 13: The process of generalisation by increasing the spacing (1) or decreasing the resolution (2)

It is a common thought that in raster models, the size of the resolution always equals the spacing between the cells ($r = s$). This does not have to be the case. In block methods the r is smaller than the s and in focal methods the case is vice versa. In this study the original resolution of a cell is 25 metres and the original spacing is 25 metres ($r = s$), but after generalisation (for example generalising the cells by increasing the spacing between the cells 5x) the spacing is increased to 125 metres, so $r < s$ (see Figure 13). If the digital image is represented correctly, every grey cell (grey representing the class 'grass' in this case) should be surrounded by a halo of two cells wide in all eight directions (24 cells) of colourless cells, colourless representing no information (McBratney, 2001). This is rarely seen, instead GIS programs tend to fill up the empty space with the information in the middle cell. This is a misrepresentation of the data and often causes confusion to the user. Therefore, in the case whereby s is increased while r is held constant, only a fraction of geo-information is actually carried through in the process while a lot of information is lost. In Figure 13.1, $1/25^{th}$ of the geo-information in the block is later available for spatio-temporal analysis but the rest is lost in the generalisation process.

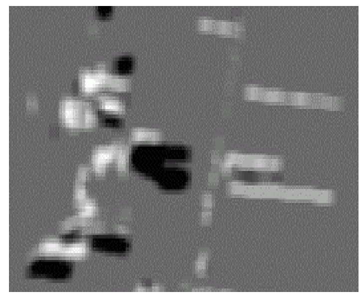
Alternatively, the r can be bigger than the s ($r > s$), this is the case in focal methods. Here, the cell size which is normally represented by r, is now represented by s, the spacing between the information. Only the middle part of the resolution can be seen. In Figure 13.2, the resolution of the bold bordered cell is 125 metres and the spacing is 25 metres . In this situation, geo-information available in the original data source is maintained through the generalisation process. This is useful information when spatio-temporal analysis is done because the variation can be analysed at cell-level. This particular generalisation process (holding s constant and increasing r) is most suitable for data that are used for spatio-temporal analysis with a low-level widespread change such as detecting spatio-temporal variation of 'grass'. The most suitable generalisation process for the purpose of this particular temporal analysis is therefore one that keeps as much information as possible. It is obvious that for this type of spatio-temporal analysis it is better to use focal methods rather than block methods in order to keep the geo-information available in the original data. One might argue that focal methods result in overlapping information, however although the overlapping blocks are strongly correlated with each other, there is still a lot of information present (and hardly any noise) (McBratney, 2001).

Figure 14 displays the images of the observed spatio-temporal variations obtained from the original data sets and generalised data sets. The focal method produces a much more similar image to the spatio-temporal variation obtained from the original data sets than when generalised by the block method. This suggests that generalisation by focal methods (increasing r) maintains more original geo-information than generalisation by block methods (increasing s).

(o_1-o_2) between original data sets



(o_1-o_2) between data sets generalised by the mean block method



(o_1-o_2) between data sets generalised by the mean focal method

Figure 14: Section of spatio-temporal maps showing the difference between the observed spatio-temporal variation obtained from the original data set and the observed spatio-temporal variation after generalising the data set 5x

(iii) The effect of the decision rule on the observed spatio-temporal variation

Apart from the relation between the *r* and the *s*, the decision rule used to merge geo-information is very important in a space-time dimension. The majority decision rule results in a loss of original geo-information (Figure 4.1 and 4.2). The percentage of observed spatio-temporal variation obtained by using the mean and the weighted average method follow a more similar decreasing trend compared to the percentage variation obtained from the majority method, which shows a deviation from this trend (Figure 15).

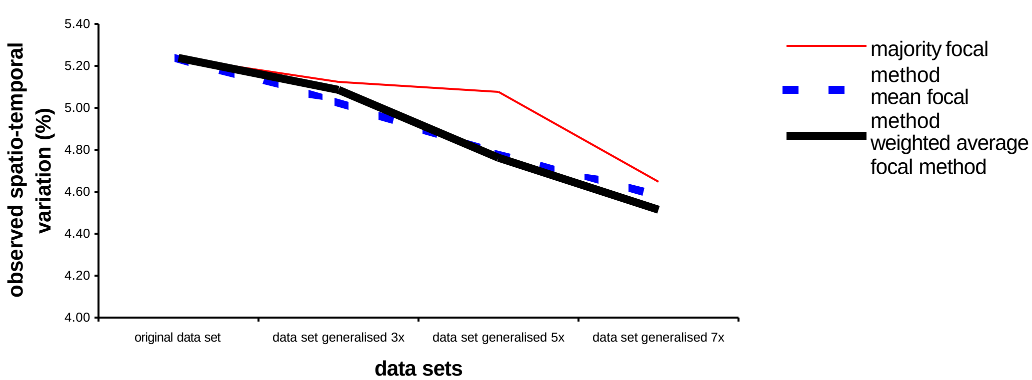
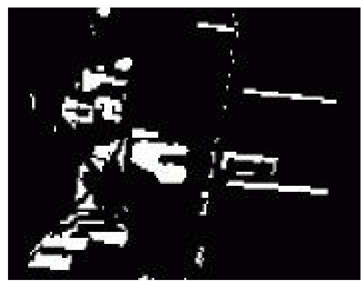
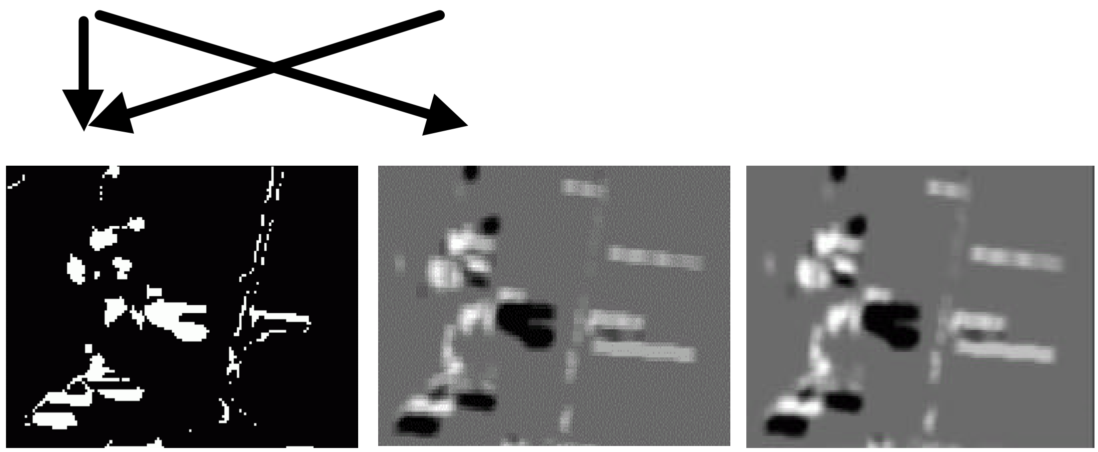


Figure 15: Relation between the observed spatio-temporal variation and generalisation with the majority focal, mean focal and weighted average focal methods

When geo-information from a block of 9 cells is generalised to 1 cell by the majority rule, not all geo-information is carried through to the new generalised block. Instead, a choice has to be made to decide which class, ‘grass’ or ‘not grass’ becomes the new geo-information in the generalised cell (see Figure 4.1 and 4.2). Once this choice has been made, other geo-information is eliminated. Less geo-information is lost when the mean or the weighted average is calculated as the value of the new cell takes into account all geo-information in its neighbourhood by presenting the value as a fraction of ‘grass’ or as a weighted average (Figures 4.3, 4.4 and 5). The decreasing trend of the percentage of spatio-temporal variation that follows the majority rule in its generalisation process (from 5.24% to 4.65% using the focal method and from 5.24% to 4.20% using the block method) is largely due to the loss of geo-information (Figure 15). This loss of geo-information due to the majority rule can be clearly seen in the images below when compared to the observed spatio-temporal variation from the original data sets (Figure 16). The observed spatio-temporal variation obtained after generalisation using the mean and the weighted average decision rules are more similar to the observed spatio-temporal variation from the original data sets.



(o_1-o_2) between original data sets



(o_1-o_2) between data sets generalised by the majority focal method

(o_1-o_2) between data sets generalised by the mean focal method

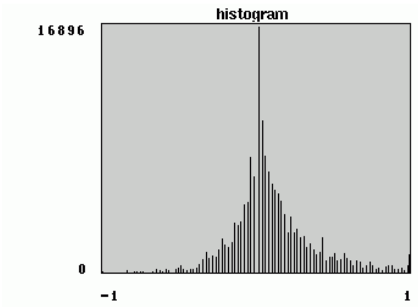
(o_1-o_2) between data sets generalised by the weighted average focal method

Figure 16: Section of the spatio-temporal maps showing the difference between the observed spatio-temporal variation obtained from the original data sets and the observed spatio-temporal variation obtained after the generalising the data sets 5x

Detecting finer observed spatio-temporal variation (o_1-o_2)

Block methods and the majority decision rule clearly result in a loss of geo-information at cell-level. The decrease in observed spatio-temporal variation when the mean and the weighted average focal methods were applied to generalise, however, is more due to the filtration of noise than the loss of geo-information. These methods have used all available information in the original data sets and have presented the new cells with values that reflect the geo-information of the new resolution. They can therefore be considered as the most appropriate of the proposed generalisation methods for detecting low-level widespread change such as the monitoring of the change of land-use over time.

Burrough and McDonnell (1998) emphasised the importance of recording variations that occur at cell-level when dealing with both the space and time dimensions. Calculating the weighted average using Block Kriging may therefore be a slight improvement to the mean method as it uses geo-information from outside the block to estimate the average, the new value of the generalised cell. This method provides more detailed spatio-temporal variations as more cells are involved in its calculation. Due to the filtering of noise and the representation of more detailed geo-information, finer observed spatio-temporal variations may be detected at cell-level (Figure 17). This is indeed very useful for spatio-temporal analyses, because these finer variations can be localised.



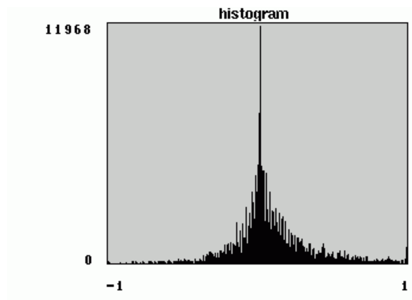


Figure 17: Histograms showing the observed spatio-temporal variations at cell-level; (1) spatio-temporal cell values generalised with the mean focal method (stdev=0.285) (2) spatio-temporal cell values generalised with the weighted average focal method (stdev=0.266)

The program VESPER (used to calculate the weighted averages) also provides information on the uncertainty of the new estimate, which none of the other methods do, although they can be calculated. In this method the spatial correlation between the cells are taken into account which is not the case when the mean method is applied. The mean method simply takes the average of the values of the cells within the block without taking into account the values surrounding the block. The main strengths of kriging are in the statistical quality of its predictions (e.g. unbiasedness) and in the ability to predict the spatial distribution of uncertainty (Mitas and Mitasova, 1999). The variance calculated by VESPER is a function of the spatial variation of the class 'grass', and the configuration of the cell to be estimated. With this information, statistically significant variations may be calculated at cell-level. Due to time limitations this has not been done. Nevertheless, this method potentially provides a more detailed representation of the observed spatio-temporal variation than any of the other proposed methods. It needs to be highlighted that this method of generalisation is specifically useful for detecting low-level widespread change, it is not necessarily the most suitable generalisation method for other spatio-temporal analyses.

The optimal cell size for spatio-temporal analyses

The decreasing trend of the observed spatio-temporal variation when cells are generalised by increasing *r* is related to the trade-off between the loss of geo-information due to averaging over greater areas and the increased information due to filtering out of random noise. As the resolution is decreased a point will be reached where the loss of geo-information is greater than the filtration of noise. The optimal cell size will therefore be at a resolution where the loss of geo-information is minimal and the gain of temporal resolution is maximal so that the filtering of noise is maximised. Using Block Kriging the optimal cell size can be calculated. The optimal value (weighted average) of the cell can be calculated when the uncertainty of each value is known. As the program VESPER provides this information, the resolution where the optimal averaging occurs (with minimum uncertainty) can be calculated. However, due to time limitations this was not done.

6. CONCLUSION

Generalisation in GIS results in a less precise representation of the original data source. However this does not necessarily imply a degraded representation of reality. It enhances certain aspects and reduces others for the purpose of its application. In this study our purpose is to obtain the spatio-temporal variation of the land-use 'grass' that is most representative of the actual change in the real world. In spatio-temporal analyses, the dimensions of space and time are combined. It is in such a situation that cells should contain detailed and representative geo-information in order to obtain observed spatio-temporal variation (o_1-o_2) representative of the true variation (m_1-m_2).

Results have shown that generalisation does influence the state of noise in a space-time dimension. Noise filtration does occur by a generalisation process. Generalising data for spatio-temporal analyses is therefore beneficial to the quality of the results. An increase in the *r* or the *s* means that averaging occurs over a larger area which results in a decrease of the variance due to noise. Due to this filtration of noise after the generalisation process, the observed spatio-temporal variation is probably more representative of the true spatio-temporal change in the real world.

Not only the filtering of noise occurs when cells are spatially generalised, loss of geo-information may also occur depending on how generalisation is processed. Generalisation whereby *r* is increased seems to be more appropriate for detecting low-level wide-spread change. Geo-information available in the original data source is maintained through the generalisation process unlike the case when *s* is increased. Increasing *s* results in a loss of geo-information as only a portion of the original geo-information is available after the process.

The majority decision rule cannot be recommended when the purpose of generalisation is to detect change. By using the majority rule not all geo-information is carried through to the new generalised cell. Instead, a choice has to be made to decide which value becomes the new geo-information in the generalised cell. More appropriate decision rules are when the mean or the weighted average of a generalised cell is calculated. Calculating the weighted average using Block Kriging and increasing *r* (while holding *s* constant) in the generalisation process may be the most suitable method for detecting spatio-temporal variation. During the generalisation process, geo-information from outside the block is used to estimate the new value of the generalised cell. Due to the filtering of noise and the representation of more detailed geo-information, finer observed spatio-temporal variations may be detected at cell-level which are more representative of the true change in the real world.

Overall, when combining space and time in geographical analyses, data are required to have the properties of scientific observations. They must be relevant surrogates for the environment they represent. Unfortunately, this is not the case with most remote-sensed data, due to impediments during the identification process. The presence of noise is therefore impossible to be completely eliminated. However, filtering of noise is proven to be possible by means of spatial generalisation, which may be one way to bridge the divide between the digital and the real worlds.

7. FURTHER STUDIES

Due to the absence of a reference data set the results cannot confirm conclusively that spatial generalisation using the weighted-average focal function is the most useful method out of all proposed methods for detecting land-use change. A reference data set is needed to see whether the observed spatio-temporal variation after this generalisation method is closest to the true change in the real world compared to other methods. Therefore a similar study with a reference data set is recommended.

Burrough and McDonnell (1998) expressed the importance of the choice of the cell size because variations that occur within a space-time dimension will not be registered by either the data or the process. Using Block Kriging the optimal cell size can be calculated. The optimal value (weighted average) of the cell can be calculated when the uncertainty of each value is known. As the program VESPER provides this information, the resolution where the optimal averaging occurs (with minimum uncertainty) can be calculated. Due to time limitations, the optimal *r* was not calculated, however this is strongly recommended for further studies.

The poor quality of spatio-temporal analyses that use classified data is partly due to the uncertainty of the characteristics of the environmental variables. Therefore, generalising environmental variables might not be the best way to get closer to the detection of true spatio-temporal change. Some variables in the spatial data sets are uncertain as one cannot be sure whether these are correctly mapped. When combining a space-time dimension it may be more useful to generalise before data are classified in order to prevent additional noise due to classification. For example, performing Block kriging on the individual bands of remotely-sensed data is recommended, prior to classification and temporal analysis.

Acknowledgement

This paper is a follow-up of the ORGISM project at the Centre for Geo-Information in Wageningen, lead by Kees de Zeeuw. Many thanks to him and Gerard Hazeu for providing the data, recommending this research and their guidance at the start of this study. Thanks to Sytse de Bruin for constructive discussions. Many thanks also to Alex McBratney for his intuitive ideas, helpful advice and continuing interest throughout the study.

REFERENCES

Al-Taha, K.K. and Barrera, R. (1990) Temporal data and GIS: an overview. Proceedings of the GIS/LIS’90 Conference, 1, 244-254

Bundy, G.L., C.B. Jones, C.B. and E. Furse (1995) *Holistic generalisation of large-scale cartographic data* In: *GIS and Generalisation - Methodology and Practice.* Müller, Lagrange and Weibel (Eds), Taylor & Francis, pp. 106 - 118

Burgess, T.M. and R. Webster (1980) Optimal Interpolation and Isarithmic Mapping of Soil Properties: I. The Semi-Variogram and Punctual Kriging. *Journal of Soil Science*, Vol 31, pp. 315-331

Burrough, P.A. and R.A. McDonnell (1998) *Principles of Geographical Information Systems.*. Oxford University Press

Buttenfield, B.P. (1995) Object-oriented map generalisation: modelling and cartographic considerations. In: *GIS and Generalisation - Methodology and Practice.* Müller, Lagrange and Weibel (Eds), Taylor & Francis, pp. 91 - 104

Curran, P.J., E.J. Milton, P.M. Atkinson and G.M. Foody (1998) Remote Sensing: From Data to Understanding. In: *Geocomputation A Primer*. P.A. Longley, S.M. Brooks, R. McDonnell, B. Macmillan (Eds). John Wiley & Sons Inc. pp.33-59

Fabbri A.G. and C.J.F. Chung (1996) Predictive spatial data analysis in geosciences. In: *Spatial analytical perspectives on GIS.* Fischer M., H.J. Scholten and D. Unwin (Eds). Taylor& Francis, London, United Kingdom

João E.M. (1995) The importance of quantifying the effects of generalisation. In: *GIS and Generalisation - Methodology and Practice.* Müller, Lagrange and Weibel (Eds), Taylor & Francis, pp. 183 - 192

Langran, G. (1992) *Time in Geographic Information Systems.* Taylor & Francis, London, United Kingdom

McBratney, A.B. (2001) *Resolution, spacing & extent - some thoughts on digital maps.* Department of Agricultural Chemistry & Soil Science, The University of Sydney. (personal communication)

McMaster R.B. and Shea K.S. (1992) *Generalisation in digital cartography.* Washington DC. Association of American Geographers

Minasny,B., McBratney, A.B., and Whelan, B.M. (1999) *VESPER version 1.0.* Australian Centre for Precision Agriculture, McMillan Building A05, The University of Sydney, NSW 2006 (http://www.usyd.edu.au/su/agric/acpa)

Mitas, L. and H. Mitasova, (1999) Spatial interpolation. In: *Geographic Information Systems*. P.A. Longley, M.F. Goodchild, D.J. Maguire, and D.W. Rhind.(Eds). John Wiley & Sons Inc. Vol. 1, 2nd Ed., pp.481-492

Müller, J.C., R. Weibel, J.P. Lagrange and F. Salgé (1995) Generalisation: state of the art and issues. In: *GIS and Generalisation - Methodology and Practice.* Müller, Lagrange and Weibel (Eds), Taylor & Francis, pp. 3-17

Noordman, E., H.A.M. Thunissen & H. Kramer (1997) *Vervaardiging en nauwkeurigheid van het LGN2-grondgebruiksbestand*; Achtergrondinformatie bij gebruik van het bestand. Wageningen, DLO-Staring Centrum, Rapport 515. (Dutch)

Oosterom van, P. (1995) The GAP-tree, an approach to ‘on-the-fly’ map generalisation of an area partitioning. In: *GIS and Generalisation - Methodology and Practice.* Müller, Lagrange and Weibel (Eds), Taylor & Francis, pp. 120 –132

Peuquet, D. (1994) It’s about time: a conceptual framework for the representation of temporal dynamics in geographic information systems. Annals of the Association of American Geographers, 84(3), 441-461.

Roughani, S. and H. Wackernagel (1990) Multivariate Geostatistical Approach to Space-Time Data Analysis. In: *Water Resources Research*, Vol. 26, no. 4, pp. 585-591

Sonneveld, M.P.W. J.A.T. van den Brink, C.J. de Zeeuw and A.K. Bregt (2000) Noise reduction in GIS-based land use monitoring. *Proceedings 9th inernational symposium on spatial data handling*, Beijing, China. Section 5a, pp 44-56

Schoorl, J.M., A. Veldkamp and L.O. Fresco (1997) The conversion of Land use and its Effects. A regression based model applied to Costa Rica (Pascal version 1.2.) *Quantitative approaches in system analysis* no. 8, AB-DLO, Wageningen, The Netherlands

Star, J. and J. Estes (1990) *Geographic Information Systems An Introduction.* Prentice-Hall Inc. New Jersey

Wachowicz, M. (1999) Object Oriented Design for Temporal GIS. Taylor & Francis

Wachowicz, M. (2000) The Role of Geographic Visualisation and Knowledge Discovery in Spatio-Temporal Data Modelling. In: *Proceeding Time in GIS* workshop…

Weibel, R. and G. Dutton (1999) Generalising spatial data and dealing with multiple representations. In: *Geographic Information Systems.* P.A. Longley, M.F. Goodchild, D.J. Maguire, and D.W. Rhind.(Eds). John Wiley & Sons Inc. Vol. 1, 2nd Ed., pp. 125-155

Zeeuw, C J. de and G.W. Hazeu, G. (Eds) (2001) *Monitoring land use changes using geo-information;. Possibilities, methods and adapted techniques,.* Centre for Geo-information, Wageningen, Alterra, Green World Research