GEOGRAPHICAL INFORMATION PROCESSING

Towards transparent statistical mapping

By

Jan Ketil Rød

Dr.Polit thesis 2002
Department of Geography
Faculty of Social Science and Technology Management
Norwegian University of Science and Technology, NTNU
Trondheim, Norway

ISBN 82-471-5091-3

Acknowledgements

This Ph.D. project would not have been started without Axel Baudouin's enthusiasm which initiated my interest in graphic semiology and geographical information processing. Thank you for your support and refreshing views when disagreeing with me. I indeed benefited from your network when I came to Paris to take a DEA in geography. Especially Lena Sanders was helpful and friendly to me allowing me to use Equipe Paris' library and computer room at 13 rue du Four, the place I met Jean-Marc Orhan. We had similar projects and interests which motivated me significantly in this early phase of the project. Although you teased me too often for my very bad French, I indeed enjoyed your dinners. Thanks also for the help you gave me together with Névina Mintcheva in writing an article in French. In 1998, Jean Marc organised Colloque 30 ans de sémiologie graphique, where I for the first time met Ferjan Ormeling. Meeting Ferjan was so significant for me that I soon after asked him if I could come to Utrecht. Instituut voor Kartografie was certainly the place where I was most productive during the whole Ph.D. period, much due to the atmosphere created by staff and students genuinely interested in maps and mapping and Ferjan's critical but inspiring comments on the drafts for papers I wrote. Consequently, I had to ask you Ferjan to be my subsidiary supervisor. Ferjan's most *gekke en gezellige* student was Julius Spit which soon became a very good friend of me and who helped me together with Diana van der Meer in writing an article in Dutch. Thanks to Gunnhild Setten, Sissel Furuseth for their comments on the superstructure and Jon Buscall to correct my English. Thank you!

From Jon Terje Kommandantvoll I have learned a lot of programming. The fact that Jon Terje is highly skilled, shows a remarkable pride in his work as well as obviously regarding me as a good friend had given me the confidence that if I asked him to help me with a problem, the problem would be solved, recently by leading me out of the *.dll hell when making a setup for GIB.

I am also grateful for having known the staff associated with Equipe Paris, CNRS, Instituut voor Kartografie, Utrecht University and especially Geografisk Institutt, NTNU.

… and Diana, you are wonderful!

Content Part I: Superstructure

Content Part II: Articles

Content Part III: Developed software for statistical mapping

Part I

Superstructure

1 Introduction

This thesis consists of eight components: one computer program named GIB, six articles and this superstructure. The superstructure aims to show how the other seven components relate to each other and how they, together with this superstructure, constitute the thesis. A model of these relations is found in Figure 1.1.

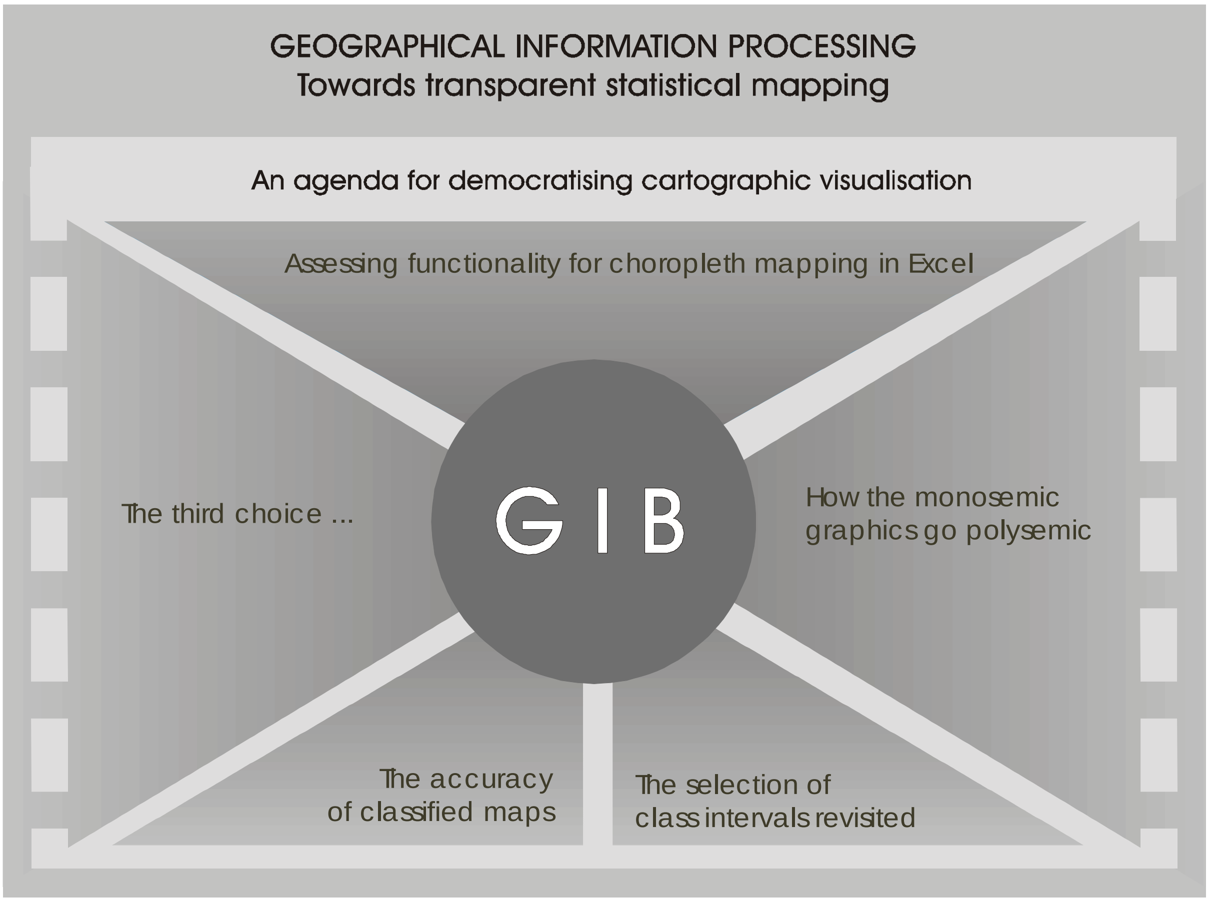


Figure 1.1: The components of the Ph.D. project Geographical information processing – Towards transparent statistical mapping.

Two of the articles can be viewed upon as motivations for developing GIB, that is *Assessing functionality for choropleth mapping in Excel* (Rød, 1999) and *An agenda for democratising cartographic visualisation* (Rød et al., 2001). As illustrated in Figure 1.1, the latter of these two articles encompasses the other articles and GIB. Cartographic visualisation and the agenda of making it available for everybody is an ongoing theme throughout this dissertation. Two of the articles, *The third choice …* (Rød, 1998) and *How the monosemic graphics go polysemic* (Rød, 2001a), are theoretical and interconnected by the superstructure, which is indicated in Figure 1.1 with openings between the superstructure and these two articles. The last two articles, *The accuracy of classified maps* (Rød, 2000) and *The selection of class intervals revisited* (Rød, VI), are based on results from the development and use of GIB.

2 Statistical Mapping

Statistical mapping is of importance in all the thesis' articles. Statistical maps are also the type of maps the GIB mapping package generates (i.e. choropleth maps, dot density maps, proportional point symbol maps and chorochromatic maps) from a flat file of units and variables stored in a spreadsheet.

2.1 Geographical Information Processing

The mapping package is called GIB, which is an abbreviation from *geografisk informasjonsbehandling*, the Norwegian term for *geographic information processing*. Bertin (1977/1981) uses a similar term: *graphic information processing*. Information is an essential key word in this term. Unfortunately, the term *information* is in general used interchangeably with the term *data*. Data is here understood to be the storage of numerical or non numerical values, whilst information has to do with knowledge. According to Bertin, information 'is the reply to a question' (Bertin, 1981: 11) and graphic information processing is the particular methods of using graphics in order to discover the answers (Bertin, 1981: 16). With graphics, Bertin includes diagrams, networks and maps (Bertin, 1967/1983). 'Graphics is a very simple and efficacious sign system which anyone can put to use' (Bertin, 1981: 16): it 'involves utilizing the properties of the plane to make relationships of resemblance, order or proportions among given sets appear' (Bertin, 1981: 176). Graphic information processing is central to this thesis, but the term is slightly altered by the prefix *geo*. This is done to emphasise an affiliation to geography. Map users' interaction with statistical maps is here viewed as a geographical information-processing problem, the particular methods of using these sorts of maps to obtain new geographical knowledge.

2.2 Statistical mapping versus topographic mapping

Statistical maps might be univariate (one variable), bivariate (two variables) or multivariate (several variables). The key emphasis is put here on univariate mapping. A univariate map portrays a single variable for census tracts, municipalities, counties or other areal units. It typically describes spatial variation with (i) shades of grey ordered according to a darker-means-more scheme (choropleth map), (ii) dots ordered according to a more-dots-means-higher-density scheme (dot density maps), (iii) point symbols ordered according to a larger-means-more scheme (proportional point symbol map), or (iv) colours ordered according to an "equal colour = equal value" and "different colour = different value" scheme (chorochromatic map). While topographic mapping portrays perceptible physical landscape features, which can be seen both by the mapmaker and the map-reader, statistical mapping portray abstracted attributes without a basis in direct observation. 'Abstract statistical phenomena are unlike tangible features of the physical environment in that they cannot be seen. This means that the mapmaker and the map-reader perceive them by a rational rather than a sensory process' (Jenks and Caspall, 1971: 218-219).

Underpinning topographic mapping is a correspondence, a positive structural likeness, between the real (geographical phenomena on the earth's surface) and its graphical representation on the map. The Latin word *representare* means "to show" or "to give a picture of", which is the meaning adopted here. The graphical representation may take different dimensional characteristics traditionally appearing on the geometric primitives of a point (zero dimensions), a line (one dimension) or an area (two dimensions) to which the real phenomena correspond. For instance, a small square on the map may correspond to a house, a black line may correspond to a railway and a blue polygon may correspond to a lake. The term topographic comes from two Greek words: *topos* meaning place and *graphein* meaning write or describe. As a topographic map thus describes places, topographic map-making is both historically and contemporarily, an essential tool for way-finding (Blakemore, 1981) and they work as such, because the correspondence between the surface phenomena in the real world and their visual representations in the map makes it possible to orient the map[1].

A similar correspondence does not exist for statistical maps. One does not, for instance, "see" class boundaries in the same manner as one "sees" physical boundaries as rivers, fjords or mountain ranges and one does not "see" political electoral preferences in the same manner as land use zones like forest, fields and urban areas are "seen". The primary aim for univariate statistical mapping is to portray the geographical distribution of a particular variable, but since both the base map and the attribute values, most often, lack visual correlates, statistical maps blot out the operation that made them possible. Users of topographic maps might, by their empirical real world observations, easily identify map inaccuracies like, for instance, a missing road junction. It is difficult, if not impossible to identify corresponding inaccuracies in statistical maps. This thesis identifies the need to make the process of statistical mapping transparent and the aim is hence to contribute to transparent statistical mapping. To make statistical mapping transparent means here to increase visibility of the way choropleth maps are imbued with classification errors by means of graphical and numerical responses (See Rød, 2000, Figures 2 and 3; Rød, VI, Figures 5 and 6). These responses will function as control checks on the generated statistical maps' quality and thus hopefully an important tool for making better decisions based on classified maps. However, in seeking such a transparency, a more profound understanding is needed of the way statistical maps refer to the real world or to concepts of the world and how statistical maps transfer or construct knowledge. These issues will be profoundly discussed in the next chapter within a theoretical frame based mostly on semiotics.

NOTES

[1] A popular representation in medieval cartography is the 'T in O' maps. The known flat world was in these circular maps surrounded by the river Oceanus, the holy city of Jerusalem were situated at the centre and they reflected the threefold division of the earth among Shem, Ham, and Japheth, the sons of Noah. 'They are called T in O maps, because they were designed with the Mediterranean as the upright part of the T, the Don and Nile rivers as the crosspiece, and the whole inside a circular ocean. The farthest area that was known at all was, of course, the Orient. It became traditional to locate Paradise in the difficult-to-reach, far eastern area and to put it at the top of the map. From this practice we have derived the term to *orient* a map – that is, to turn it so that the directions indicated are understood by the reader' (Robinson et al., 1984: 26). The correspondence is here between the map and a cathedral. Cathedrals are oriented with its entrance in the west (down on the map) so the churchgoer approaches the east (up on the map) while entering.

3 Theoretical foundation

3.1 Introduction

With this superstructure, I am aiming at sketching a theoretical foundation applicable for statistical mapping mainly based on semiotics. Harley’s incentives for an ethics and a social theory for cartography are of relevance for the subject area of this thesis and, consequently, I wish to continue the debate within cartography initiated by the publication of *Deconstructing the map* (Harley, 1989) (section 3.2.). By revisiting and revitalising the debate within a ‘representational perspective’ (MacEachren, 1995: 6) on cartography at a time when map authorship is ‘being brought to the masses’ (Lang, 1995), I hope to set the agenda for a democratised cartography in a wider context. Since ‘a semiotic perspective offers a structured way to consider the interaction of the explicit and implicit meanings with which maps are imbued’ (MacEachren, 1995: 242), an abbreviated synopsis of selected issues in the field of semiotics (section 3.3.) and its application to cartography (section 3.4.) will be outlined. This will form the base for the subsequent commentaries on the reasons why Harley (1989), who was influenced by the writings of Michael Foucault (1926 – 1984) and Jacques Derrida (1932 –) on discourse and deconstruction, did not succeed in deconstructing the map (section 3.5). Ethics, defined as principles of conduct guiding the practices of an individual or professional group (McHaffie et al, 1990: 3), have always been regarded as important among professional cartographers and relates to what Harley rather ironically coined the phrase: the ‘ethic of accuracy’ (Harley, 1989: 5). Both Harley’s and other’s ethical discussions have been calls to social and cultural issues (Harley, 1990; Chrisman, 1987; Smith, 1992; Rundstrom, 1993). The ethical concern for this thesis is a disclosure of alternative cartographic designs (Monmonier, 1991) and I will albeit on different grounds, respond to Harley’s call for a cartographic ethics (section 3.6).

3.2 The Cartographic discourse

Harley argued that ‘it is possible to view cartography as a discourse’ (1989: 12). The term *discourse*, from the French: *discours*, takes several meanings usually connected to speech, language and text. Discourse analysis can thus be understood as the analysis of speech, language and text – an academic activity normally found within linguistics and literary studies. However, discourse analysis has a much broader application, a consequence of the increasingly broadened employment the term *text* has achieved. In contemporary use, text is ‘a set of signifying practices commonly associated with the written page but over the past several decades increasingly broadened to include other types of cultural production such as landscapes, maps, paintings as well as economic, political and social institutions’ (Duncan, 2000: 824-825). Once the notion of text has been expanded to include types of cultural production other than writings then the assumption is made that these productions, i.e. maps, have a text-like quality or that they are discursive. Robinson and Petchenik (1976) refuted this idea based on an ‘assumption that linguistic syntax was equivalent to syntactics’ (MacEachren, 1995: 236). While most maps do not have elements of syntax, they, and statistical maps in particular, do have syntactical elements, which is here regarded as supporting the “map as text metaphor”. However, when Harley considered the map as text, he based his claims not on the presence of linguistic elements but on the act of construction[1] (Harley, 1989: 7). The act of construction constitutes the map as a text in the way maps construct knowledge.

The way maps construct knowledge can be viewed from two interrelated senses of map use: maps that facilitate *public visual communication* directed towards presenting knowns (presentation) and those that foster *private visual thinking* directed towards revealing unknowns (exploration)[2] (DiBiase, 1990: 14). Cartographers in general and particularly during the period in the 1970s when ‘the communication paradigm’ (MacEachren, 1995) was pre-eminent have understood maps primarily as vehicles for transfer of information or as a medium for presentation. Maps used for presentation represent a form of knowledge construction, which is produced by the way the maps “transfer” particular geographic information. This corresponds to what Gregory has described as ‘the naturalizing function of discourse’ and ‘the situated character of discourse’ (Gregory, 1994: 136). The naturalising function of cartographic discourse shapes the contours of the taken for granted world: they “naturalise” and often implicitly universalise a particular view of the world. As Harley expressed it, we need to consider how maps shape ‘mental structures’ and impart ‘a sense of the places of the world’ (Harley, 1989: 13). In this way, maps may imbue or universalise a particular view of the world and, consequently, knowledge being communicated or transferred by maps cannot be neutral or free from power relations. Propaganda maps, development maps, statistical maps, and even “innocent” maps like topographic map series possess rhetorical power (see e.g. Wood, 1992). The situated character of discourses provides partial, situated knowledge: as such the discourses are characterised by a particular constellation of power and knowledge. The constellation of power and knowledge was among Harley's concerns and he exemplified, what he called *external* power, by the way monarchs, ministers, state institutions, and the Church have all initiated mapping for their own ends (Harley, 1989: 12). Harley also used the term *internal* power which is embedded in the cartographic process: '… the way maps are compiled and the categories of information selected; the way they are generalized, a set of rules for the abstraction of the landscape; the way the elements in the landscape are formed into hierarchies; and the way various rhetorical styles that all reproduce power are employed to represent the landscape' (Harley, 1989: 13). The power of the mapmaker is not generally exercised over individuals but over the knowledge of the world made available to people in general. Apparently, the ideas forwarded by Harley regarding power relations reflect a view of map users as passive receivers of messages: thus, geographical knowledge is imposed on them. These constellations of power and knowledge may change if mapmaking becomes a common activity allowing the "passive receivers of map messages" to be map authors (see Rød et al 2001).

The availability of massive amounts of geographic information, 'leads to the notion that having this information provides one with a better understanding of the world' (Curry, 1995: 78). However, 'because the availability of information is seen as being of fundamental importance to the making of decisions, those who have that information see themselves as empirically better able to make decisions than are those who are merely "other." And this means that there are features of the use of these systems that are fundamentally antidemocratic' (Curry, 1995: 79). If we are to follow the classical virtue: *the greatest good for the greatest number*, then we should 'provide everyone with the greatest opportunity to learn, to have free and open access to knowledge' (Sack, 1999: 39). Democratising cartography is, in principle, commensurable with providing the greatest number with the greatest opportunity to learn and to have free and open access to knowledge. 'Valuing an awareness of reality is a deeply held part of our human nature. It draws attention to our intellectual capacities to reason and to pursue truth. Making awareness public allows our views to be tested against others and clarifies the picture for all' (Sack, 1999: 34). Cartographers, like Ferjan Ormeling when commenting on Brian Harley's influence on modern cartography, have expressed similar views: 'Mapping activities should take place more because of public benefit and usefulness and democracy than has been the case in the past. This would enable concerned citizens to participate and share in defining their futures' (Ormeling, 1992: 65).

3.2.1 Map as text

In contrast to the communication paradigm, a representational perspective on cartography 'begins with an assumption that the process of representation results in knowledge that did not exist prior to that representation; thus mapping and map use are processes of knowledge construction rather than transfer' (MacEachren, 1995:459). In the realm of viewing map use as knowledge *construction* the 'map as text' metaphor may be more appropriate than the 'map as mirror' metaphor. Harley criticised the use of the 'map as mirror' metaphor in the context of topographic mapping (Harley, 1989: 4) and suggested that it ought to be replaced with the 'map as text' metaphor (Harley, 1989: 7). Although the term 'mirror' is not part of the 321 definitions of the word 'map' investigated by Andrews (1996), it is used in statements like 'the map mirrors the world' (Muehrcke, 1978: 298) and 'maps that reflect reality' (Papp-Váry, 1989: 104). Besides, the old Latin name for atlases: *speculum orbis terrarum* indicate that atlases were not representing the earth, they were mirroring it. I believe that relatively few contemporary cartographers would support any of the two binary extremes of a map metaphor, but rather take an intermediate position. However, an understanding towards the 'map as text' metaphor might be more relevant in the context of statistical mapping than for topographic mapping for two reasons:

(1) The correspondence between real surface objects and their graphic representation in the map is not present in the statistical map in the same manner as with topographic maps and thus emphasises the notion of construction.
(2) In making the tools and data of map authorship widely accessible, personal computers and mapping software have fostered a democratisation of cartography (see e.g., Lang, 1995), which is visible first of all in the field of statistical mapping.

The democratisation of cartographic visualisation may imply two frightening prospects: (1) an increased use of statistical maps for persuasive communication, designed consciously to support a particular hypothesis, viewpoint or political agenda; and (2) an increased production of poorly designed maps. In this thesis I take a somewhat naïve point of view believing in two other beneficial results arising from the democratisation of statistical mapping. If packages for statistical mapping are equipped with reasonable default options and possibilities for users to make the "right" choices regarding map design, I believe the production of poorly designed maps will be avoided. In addition, it is possible that maps as symbols of authority will diminish and their reputation as objective, value-free images of the world will decline or even vanish. Providing everybody with the opportunity to produce their own statistical maps may well replace (if it is still there) the view of the map as a mirror of the world with a view of the map as an argument equivalent to a textual or verbal assertion (Rød, 1998: 36).

3.2.2 The crisis of representation

Foucault’s basic discursive entity is the “énoncé,” or statement (Belyea, 1992: 4) and as his field of study ‘Foucault defines the organization of énoncés, the “archive”’ (Belyea, 1992: 4). Énoncés are grouped ‘according to their functions within various discursive practices. These discursive formations are regulated in turn by a mechanism of possibilities and restraints’ (Belyea, 1992: 4-5). A perspective on discourse thus, is to view it as a system setting the conditions for knowledge – that is which rules govern the production or representation of knowledge. Applied to cartography, discourse is seen as ‘a system which provides a set of rules for the representation of knowledge embodied in the images we define as maps and atlases’ (Harley, 1989: 12). Harley focused on two sets of rules that have formed the history of cartography or that have directed the advancements within cartography: ‘One set may be defined as governing the technical production of maps. […] The other set relates to the cultural production of maps’ (Harley, 1989: 4). While the former set of rules can be defined within and by using concepts from science, the latter set of cartographic rules needs to be understood in a broader historical context than merely scientific procedures or technique. Thus, instead of being solely occupied with the technical aspects of cartography where maps are seen as neutral representations of nature, Harley suggested that cartographers should accept maps as socially constructed images (Harley, 1990: 6). The aspects of the cultural production of maps have not, according to Harley, been adopted within the cartographic society. On the contrary, Harley claimed, cartographers have tried to distance themselves from those aspects[3]. As a consequence, the distance or the gap, between these two dimensions of mapping constitutes what Harley denotes as a ‘crisis of representation’ (Harley, 1990: 7, 9). According to Harley, the combination of these two dimensions of mapping will only be possible with an epistemology rooted in social science rather than positivism. Harley used expressions like ‘normal science’ (1989: 4), ‘crisis’ (1990: 7, 9) and ‘epistemological shift’ (1989, 1) thus addressing, in a Kuhnian sense, a scientific revolution for cartography. According to Kuhn, periods of scientific revolution commence when the ruling paradigm cannot account for unbearable anomalies. The *normal science* is then in a state of *crisis* before it becomes replaced by a new paradigm, which is incommensurable with the former (Kuhn, 1962). Like MacEachren, I consider the directions pointed out by Harley to be refreshing, but his ‘apparent insistence on a wholesale replacement of one limiting approach to cartography with another is not’ (MacEachren, 1995: 10). However, such a wholesale replacement of one approach to cartography with another is not necessary, since ‘a “traditional” cartographic approach to map meaning and a “critical social theory” approach to map meaning are not incompatible’ (MacEachren, 1995: 351).

Harley looked for a social theory in the work – or rather from commentaries on their works (Belyea, 1992: 1) – of Foucault and Derrida. However, Belyea concludes, ‘neither Derrida nor Foucault provides the “social theory” Harley would like to rely on’ (Belyea, 1992: 7). According to Belyea, Harley merely added a ‘socio-political dimension to the “reality” which maps are usually said to represent’ (Belyea, 1992: 1). Harley’s inquiry into the work of Derrida and Foucault ‘should have produced even more disturbing results. Harley’s failure to push the cartographic application of Derrida and Foucault’s arguments to their logical, radical conclusion is due in part to his imperfect reception of their ideas, and in part to the fixity of his own views’ (Belyea, 1992: 1). Harley had, according to Belyea, an imperfect understanding of deconstruction and she doubts whether he really did what he claimed to do:

> I shall specifically use a deconstructionist tactic to break the assumed link between reality and representation which has dominated cartographic thinking (Harley, 1989: 2).

Harley refused to consider the map as a “mirror of nature” and found “text” a better metaphor for the map. He wanted to draw upon Derrida’s central positions on deconstruction when he wanted to break the assumed link between reality and map. Derrida builds on and contributes to an extensive commentary on Ferdinand de Saussure’s theory of language (Derrida, 1997), especially the arbitrary nature of the sign, which is a prerequisite for deconstruction. Although Harley seems to have the arbitrary nature of the sign in mind: ‘It would be naive to think that either a traditional map or the latest Geographical Information System will ever approach a unity of signifier and signified’ (Harley, 1990: 8), Belyea claims that ‘the connection between cartographic signs and the world they “represent” is more complex than Harley acknowledges’ (Belyea, 1992: 4):

> Once the arbitrary nature of the sign is acknowledged, and the links of progressive degradation and externalization from thought to speech to writing are “deconstructed”, then it is obviously that *no* sign, whether verbal or graphic, literal or figurative, can be “properly” or “naturally” identified with objects or

> places or ideas outside language […] both verbal and graphic signs are purely conventional designations. Moreover, if signs do not directly indicate outside or preverbal truth, but point instead to other signs – […] – language and other sign systems must be said to function not by representing thought and nature, but by establishing and adjusting purely arbitrary relationships within each system (Belyea, 1992: 4, originally italic).

Belyea is not pleased with Harley's acceptance of the orthodox definition of maps as 'graphic representations of the world' (Belyea, 1992: 1), which seems to be the reason why she believes his inquiry did not produce even more disturbing results from his "deconstruction" of the map. In the next sections, taking a semiotic perspective, I will propose that Harley could not escape the "orthodox" conception on map representation because the cartographic sign cannot be arbitrary in a Saussurean sense.

3.3 The general study of signs

As outlined in Rød (2001a), there are two dominant semiological traditions, one European, influenced by Ferdinand de Saussure (1857 – 1913), called semiology and one North American influenced by C.S. Peirce (1839 – 1914), called semiotics. Semiology and semiotics are notions generally used by French and Anglo-American writers respectively, but they refer to the same discipline: the general study of signs. The two traditions differ in their general model of sign referred to as *dyadic* and *triadic* models, alluding to the number of elements identified in their sign relationships. In Saussure's dyadic sign model the sign is the union of the two sides that constitute it: a signified and a signifier. The fact that signified and signifier are both mental entities and independent of any external object in Saussure's theory of the sign, is the most apparent difference from Peirce's sign model. The Saussurean terms signified and signifier can be comprehended as equivalent with the Peircean terms interpretant and sign-vehicle respectively. In this superstructure, both Saussure's and Peirce's concepts will be used, the latter in closed brackets. However, only *interpretant* and *sign-vehicle* among the three Peircean categories have their counterparts in the Saussurean dyadic sign model as shown in Figure 3.1. In Peirce's triadic model the referential object, which is a term having several names (e.g. designatum) is included as a third category.

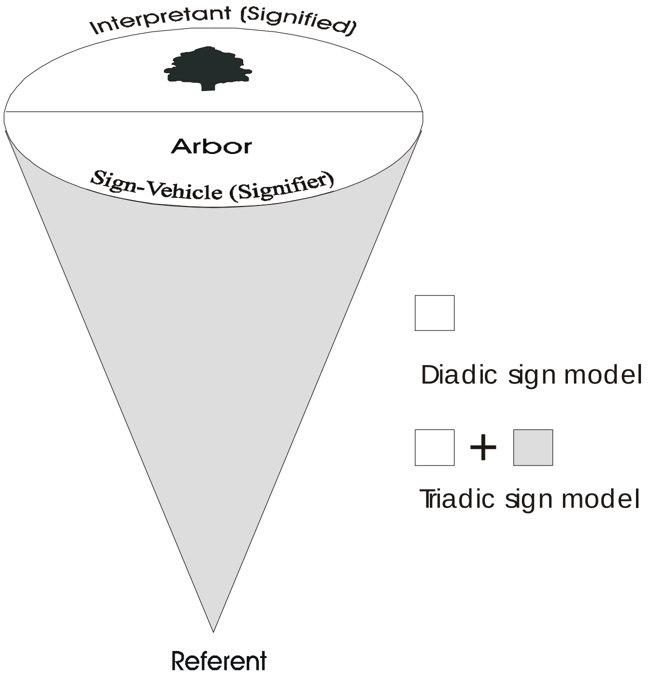


Figure 3.1: The components of a dyadic and a triadic sign model.

Saussure explicitly rejected the referential object as an element in his semiology, for 'the linguistic sign unites, not a thing and a name, but a concept [signified] and a sound-image [signifier]' (Saussure, 1974: 66). 'For Saussure, nothing existed (structurally) beyond the signifier and the signified. His semiology operated totally within the sign system. Since only a semiological system gives structure to the otherwise amorphous world, the referential object is excluded from semiotic consideration' (Nöth, 1990: 60-61). Saussure's theory of the sign had nothing to do with how sign–vehicles [signifiers] refer to real-world entities, only with how they refer to mental concepts. Applying this view to mapping, MacEachren states, 'we arrive at the conclusion that maps do not refer to the real world, but to concepts about the world' (MacEachren, 1995: 220). From the perspective of viewing the map as referring to concepts of the world, the map can be understood as a self-referential system – still a representation, not comprehended as mirroring an independent real world, but as a construction according to the knowledge possessed by individuals who work for and within various social institutions. It is in this view that Harley argued that cartography might be looked upon as a discourse as outlined above.

3.3.1 Saussure's first and second principle

According to Saussure's theory of the linguistic sign, the sign has two primordial characteristics, which Saussure called the first and the second principle: (1) the arbitrary nature of the sign and (2) the linear nature of the signifier (Saussure, 1974: 67). ‘In enunciating them I am also positing the basic principles of any study of this type’ (Saussure, 1974: 67). Thus, ‘Saussure’s semiology presented language as the analytical paradigm for all other sign systems’ (MacEachren, 1995: 217). Saussure’s first principle stated that

> the bond between the signifier and the signified is arbitrary. Since I mean by sign the whole that results from the associating of the signifier with the signified, I can simply say: *the linguistic sign is arbitrary* (Saussure, 1974: 67 – orginally italics).

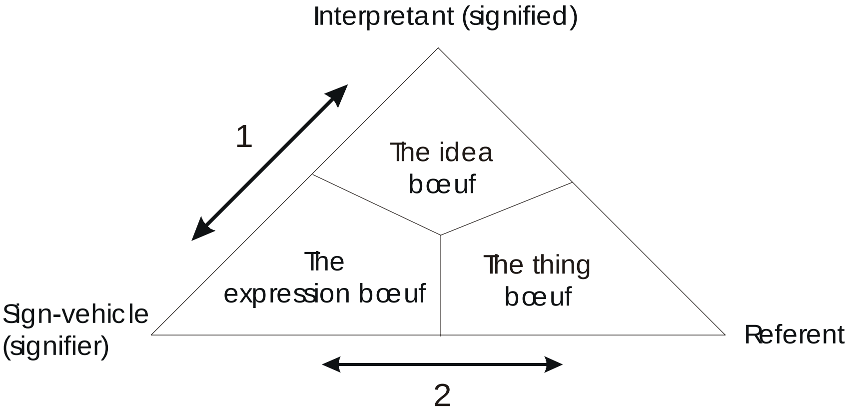


Figure 3.2: The two disputed conceptions of arbitrariness: (1) as a relationship between the signifier and its conceptual signified (sense) and (2) as a relationship between the signifier and extra linguistic reality (reference).

There has been disagreement concerning the question whether arbitrariness is a relation between the signifier and extra linguistic reality (reference) or between the signifier and its conceptual signified (sense) (Nöth, 1990: 243). These two alternative relationships constituting arbitrariness are illustrated in Figure 3.2.

Saussure’s theory of the sign strictly excludes any semiotic consideration of extra linguistic reference. The orthodox interpretation of his first principle would therefore state that arbitrariness is a matter of sense only; i.e. a relationship between the signifier and the signified (relationship 1 in Figure 3.2). The signifier and the signified are the two sides of a “psychological entity”, the sign whose relationship Saussure claimed to be unmotivated. ‘Meaning, according to Saussure, is a differential value, determined only by the structures of the language system and not by any extralinguistic reality. Since languages differ both in phonetic (*soeur* vs. *sister*) and in semantic (*mouton* vs. *sheep*[4]), the language sign is an arbitrary combination of arbitrary segments of semantic (conceptual) and phonetic substances' (Nöth, 1990: 243 – originally italics). According to Nöth, the idea of structure was fundamental to Saussure's concept of language as a system of values, an idea he expressed in a metaphor of chess: 'Saussure pointed out that only two things matter in chess: the value of the pieces according to the rules of the game, and their positions on the chessboard' (Nöth, 1990, 195). Likewise 'each linguistic term derives its value from its opposition to all the other terms' (Saussure, 1974: 88). Elsewhere Saussure has concluded: '*There are no signs, there are only differences between signs*' (Nöth, 1990, 195 – originally italics).

Saussure, when outlining the second principle on the linear nature of the signifier, made a difference between auditory and visual signifiers:

> The signifier, being auditory, is unfolded solely in time from which it gets the following characteristics: (a) it represents a span, and (b) the span is measurable in a single dimension; it is a line. […] In contrast to visual signifiers (nautical signals, etc.) which can offer simultaneous groupings in several dimensions, auditory signifiers have at their command only the dimension of time. Their elements are presented in succession; they form a chain (Saussure, 1974: 70).

In the following, it will be shown how cartographers have adopted the differentiation Saussure made between auditory and visual signifiers and how they have rejected the idea of the arbitrary nature of the cartographic sign. It will be proposed, based on several arguments that just like visual signifiers do not follow Saussure's second principle on linearity, cartographic signs do not follow Saussure's first principle on arbitrariness. The reason why cartographic signs cannot be arbitrary is found to be their referential nature and their degree of iconicity (see section 3.5). Toponomies or place names are exceptions on cartographic signs, which indeed are arbitrary. Toponomies are arbitrary because they do not refer to the "real" world, but to concepts about the world (Ormeling, 1983), e.g., naming practice like using Samaria and Judea for the occupied territory at the West Bank on Israeli maps (Rød, 2001b).

3.4 Semiology of graphics

In 1967 Jacques Bertin published *Sémiologie graphique: Les diagrammes, les réseaux, les cartes* (Bertin, 1967), which was translated into English in 1983; *Semiology of graphics: diagrams, networks, maps* (Bertin, 1983). Although Bertin uses the term semiology and many other Saussurean terms, 'no citations are given to other works which employ these terms'

(Board, 1981: 60). An explicit indication that Bertin takes Saussure's ideas as a basis for his own research is the adoption of the way Saussure differentiated between auditory and visual signifiers in his second principle, while an explicit indication on the contrary is a rejection (or avoidance) of Saussure's first principle.

Bertin compared graphics with mathematics and claimed an analogy between the two because of their monosemic nature[5] (Bertin, 1981: 178). However, he also claimed they are different based on their dependency on time. According to Bertin, graphics utilises the three dimensions of the image (x, y and the "variables rétiniennes"[6]) and it is a spatial sign system independent of time. Mathematics, on the other hand, is a linear sign system defined by time (Bertin, 1981: 178).

> Remember that written notations of music, words and mathematics are merely formulae for the memorization of fundamentally auditory systems, and that these formulae do not escape from the systems' linear and temporal character (Bertin, 1981: 178).

Bertin thus emphasized the difference by the two sign systems, auditory versus graphics in the same manner as Saussure did: written notations are linearly and time dependently perceived – graphics are immediately perceived.

> We utilize graphics to save time and consequently memory; in order to SEE, that is to perceive immediately. Accordingly, a graphic which must be READ, that is perceived over time, does not solve the problem. Moreover, we observe that such a graphic is usually not even read. The reader prefers the written text, since it generally yields a much better ratio of information received to time spent (Bertin, 1981: 179, originally capitalised).

Robinson and Petchenik (1976) adopted the same position as Bertin did regarding the immediate perception of graphics. They did so, however, in a different context as they tried to argue against a linguistic approach to cartography. They presented convincing arguments that maps have no syntax and refuted the metaphor of map as language by stating that 'the two systems, map and language are essential incompatible' (1976: 43). 'They (rightly) pointed out that individual maps have no predetermined reading sequence, and therefore no "word" order comparable to that considered under the linguistic concept of "syntax." In addition, they asserted that maps have no equivalent to "words" and are not "discursive"' (MacEachren, 1995: 236). Some of the arguments forwarded by Robinson and Petchenik were based on how Langer denoted language as verbal symbolism, which is essentially discursive. 'By reason of it [discursiveness], only thoughts which can be arranged in this peculiar order can be spoken at all; any idea which does not lend itself to this ‘projection’ is ineffable, incommunicable by means of words’ (Langer, 1951: 77 – cited from Robinson and Petchenik, 1976: 50). Langer did not denote non-discursive forms of symbolism as language. ‘The laws that govern this sort of articulation are altogether different from the laws of syntax that govern language. The most radical difference is that *visual forms are not discursive*. They do not present their constituents successively, but simultaneously’ (Langer, 1951: 50, cited from Robinson and Petchenik, 1976: 50 – originally italics).

3.4.1 Syntax versus syntactics

While linguistic syntax ‘is only the study of the rules for combining words into sentences’ (Nöth, 1990: 50), syntactics is used more broadly. The term syntactics stems from Morris’ theory of semiosis and the dimensions of semiotics. Semiosis (from Greek *sǛmeiǾsis,* observation of signs) according to Morris, involves three main factors: ‘that which acts as a sign, that which the sign refers to, and that effect on some interpreter in virtue of which the thing in question is a sign to that interpreter. These three components in semiosis may be called, respectively, the *sign vehicle*, the *designatum,* and the *interpretant*’ (Morris, 1938: 3 – quoted from Nöth, 1990: 50 – originally italics), and have their roots in the triadic sign model after Peirce. ‘From the three correlates of the triadic relation of semiosis, Morris derived three dyadic relations, which he considered to be the basis of three dimensions of semiosis and semiotics. Accordingly, *syntactics* studies the relation between a given sign vehicle and other sign-vehicles, *semantics* studies the relations between sign-vehicles and their designata, and *pragmatics* studies the relations between sign-vehicles and their interpreters’ (Nöth, 1990: 50 – originally italics). A model for these three basic dimensions of semiosis according to Morris is shown in Figure 3.3.

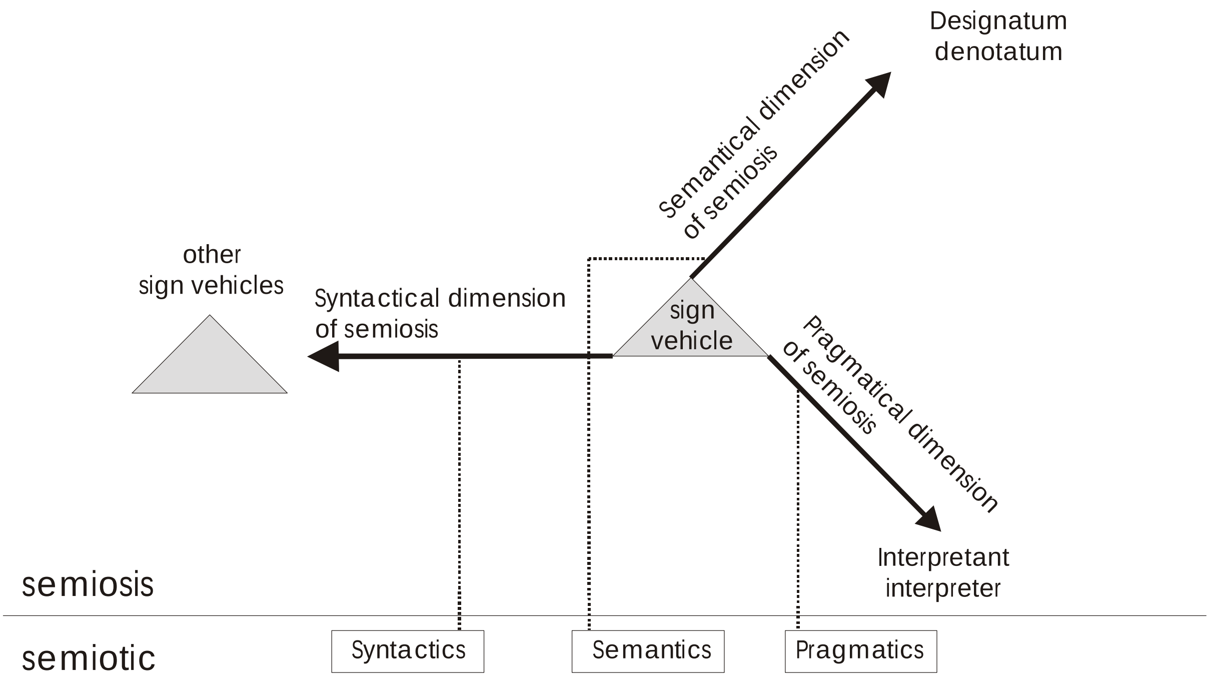


Figure 3.3. Three correlates of semiosis and three dimensions of semiotics according to Morris (redrawn from Nöth, 1990: 50).

Syntactics, the relation between a given sign-vehicle and other sign-vehicles, is the dimension of semiosis having special relevance for this thesis. Although maps might not have any syntax, as their constituents are presented simultaneoulsy, they, and statistical maps in particular, do have syntactics. ‘The point that Robinson and Petchenik (1976) missed is that while most maps do not have syntax in the narrow sense of structured reading order, they do (or should) have a carefully structured syntactics in terms of the interrelationships among signs they are composed of’ (MacEachren, 1995: 236).

3.5 Cartographic signs – are they arbitrary?

I would claim that there is an essential difference between systematising map symbolisation for topographic maps and statistical maps. In topographic mapping conventions have evolved as to how sign-vehicles [signifiers] are matched to referents: a cross on a map refers to a church, a blue line on a map refers to a river and a green area on a map refers to a forest. These are all examples of conventional symbolisation corresponding with Morris’ semantical dimension of semiosis (see Figure 3.3). The real-world referent is a critical part of the signifying relationship for a topographic map and, consequently, in topographic mapping there are several conventions of how to match the sign-vehicles to a referent. This sign vehicle ` on a topographic map would most likely be interpreted as an airport. The location the sign vehicle has on the map is essential as it designates its location on the ground. The sign vehicle for an airport would be understood differently from this sign vehicle Q, which, most likely, would be interpreted as a camping site. It is these sign-vehicles affinity to what is located on the ground that lets their meaning be derived. Not all topographic map features have the same evident affinity. Sign-vehicles for surface phenomena, as for instance contour lines have less affinity which may be the reason why they are harder to understand. Hachures or hill shading are attempts to improve the affinity between altitude and terrain variations and its cartographic representation.

The relevant signifying properties for statistical maps, on the other hand, are ‘the relationship between signs’ (Bertin, 1981: 177) and, consequently, the attempts to systematise symbolisation for statistical maps aspire to offer another set of “rules” for matching the relationship between values in a data table to relationships between sign-vehicles. These “rules” are based on Bertin’s system of visual variables, which have demonstrated ‘their appropriateness for displaying quantitative and qualitative *distinctions*’ (Buttenfield and Mackaness, 1991: 430, my italics). Bertin’s semiology of graphics (Bertin, 1967/1983) can thus be regarded as a fundamental attempt to specify a map sign syntactics in order to systematise symbolisation for statistical maps. Bertin operates with three types of distinctions: resemblance, order and proportion. Meaning is transcribed by the interrelationships between sign-vehicles, not between the sign-vehicle and the referent (designatum).

> The transcription of relationships does not utilize “signs”; it utilizes only the *relationship between signs*. It utilizes visual *variation*. Graphics denotes a resemblance *between* two things by a visual resemblance *between* two signs, the order of three things by the order of three signs (Bertin, 1981: 177 – originally italics).

According to Bertin then, it is not the sign-vehicle itself which mediates a meaning, but its interrelationships with other sign-vehicles. The interrelationships between sign-vehicles are the visual variation between them as illustrated in Figure 3.4. The visual variation between sign-vehicles showing lines in two diametrically opposite directions denotes the relationship (meaning) *difference*, the visual variation between sign-vehicles showing two different shades of grey (value variation) denotes the relationship *order* and the visual variation between sign-vehicles showing two different sizes denotes the *proportionate* relationship.

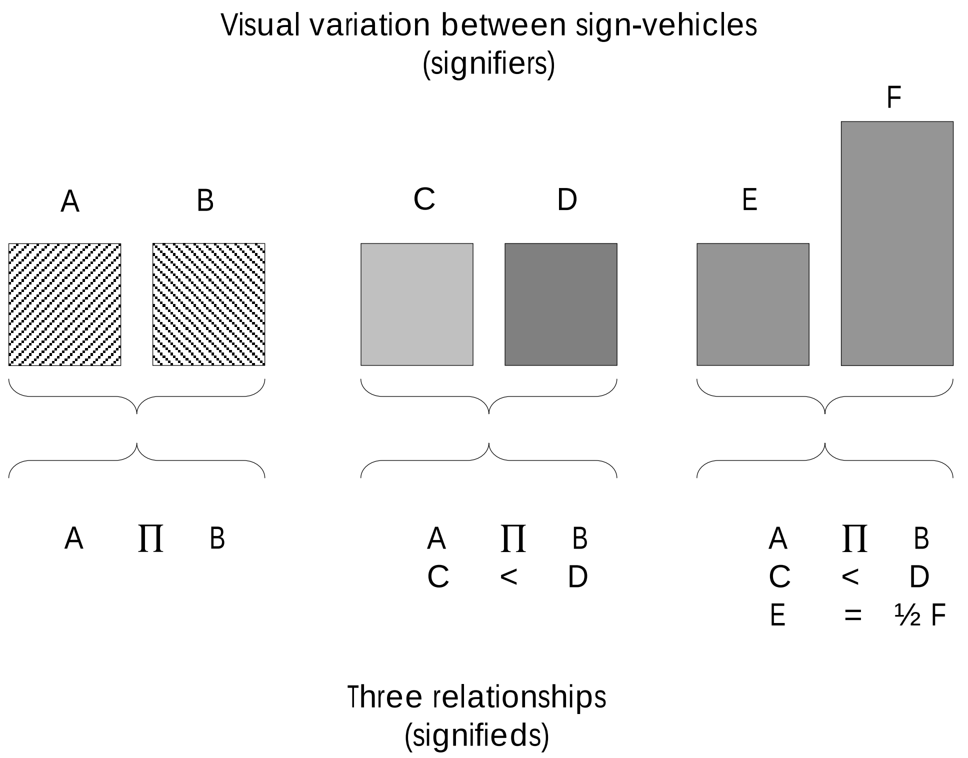


Figure 3.4. The visual variation between sign-vehicles (signifiers) denotes meaning as three relationships (signifieds) (Cf Rød, 2001a).

This seems to correspond with how Saussure regarded meaning. In Saussure's diadic sign model the referent is excluded (see Figure 3.1). Consequently, the sign-vehicle or signifier is imbued with meaning not referentially but syntactically – by the way the sign-vehicle points to other sign-vehicles. 'Meaning, according to Saussure, is a differential value, determined only by the structures of the language system and not by any extralinguistic reality' (Nöth, 1990: 243). This Saussurean premise for signifying relationships will in the following be examined for topographic and statistical maps.

3.5.1 Signifying relationship for topographic maps

To claim that the real-world referent is not a part of the signifying relationship for a topographic map would indeed create disturbing results among cartographers. These are probably the results Belyea would have expected when applying Foucault and Derrida to mapping. Then, as she expresses, also the cartographic sign system 'must be said to function not by representing thought and nature, but by establishing and adjusting purely arbitrary relationships within each system' (Belyea, 1992: 4).

Saussure specified that the term *arbitrary* means *unmotivated* (Saussure, 1974: 69). 'Unlike (iconic) signs[7] which have a "rational relationship with the thing signified," language lacks this necessary basis: "There is no reason for preferring *soeur* to *sister*". This lack of motivation thus describes the semantic dimension of the linguistic sign, i.e., the relation of sense or reference' (Nöth, 1990: 242[8] – originally italics). Cartographic sign-vehicles on topographic maps, on the other hand, are iconic and consequently, their semantic dimension is motivated. There is indeed a reason for preferring ` to Q as the sign-vehicle when an airport is to be indicated on a topographic map. These sign-vehicles have a rational relationship with the thing signified, which most people will understand independently of their mother tongue. The iconic nature of most sign-vehicles on a topographic map is the reason why these signs cannot be arbitrary in a Saussurean sense and, consequently the reason why Harley could not avoid having an "orthodox" view on cartographic representation. The implication is that any sign-vehicle having sufficient similarity to a prototype, as it might be shown in the map legend, stands for a kind of referent (e.g. any blue line on a topographic map represents a river, any green area on a topographic map represents vegetation, etc.). However, Harley did 'add a socio-political dimension to the maps traditional representation of geophysical features' (Belyea, 1992: 7), when he expressed examples of the second set of rules related to the cultural production of maps, e.g., the "rules" the mapmakers follow when they omit representations of features of the world and their use of selective toponomies.

3.5.2 Signifying relationship for statistical maps

Saussure's conceptions on meaning as a differential value (i.e. his chess example) has a striking similarity with Bertin's conceptions of the transcription of relationships:

> The transcription of relationships does not utilize "signs"; it utilizes only the *relationship between signs*. It utilizes visual *variation*. Graphics denotes a resemblance *between* two things by a visual resemblance *between* two signs, the order of three things by the order of three signs (Bertin, 1981: 177 – originally italics).

Saussure's concept of language as a system of values seems to correspond to graphics where each sign-vehicle derives its value from its opposition to other sign-vehicles. Saussure's systems of values has a correlate in Bertin's systems of relationships or the three "differential values" for graphics: resemblance, order and proportion. As shown in Rød (2001a), Bertin denotes these relationships as signifieds (Bertin, 1981: 177) and, consequently, the visual variation between visual marks can be called signifiers [sign-vehicles]. This is, however confusing when emphasising what Bertin stated that graphics do not apply signs, only relationships between them (Bertin, 1981: 177). It is more appropriate to avoid using the term signifieds – which implicitly relates the three relationships to the sign concept – but rather identifies these as Saussurean “differential values”.

I have elsewhere suggested that Bertin avoided the term “sign” as he associated it with polysemi (Rød, 2001a). According to Bertin, when ‘faced with the polysemic image, the perceptual process translates into the question: “What does such an element or collection of element signify?” and perception consists of decoding the image. The reading operation takes place **between the sign and its meaning**’ (Bertin, 1983, 2 – originally emphasis). When Bertin uses “sign” in the above citation, it refers not to the Saussurean “sign” but to the Saussurean “signifier” [sign-vehicle] and accordingly “meaning” refers to the Saussurean “signified” [interpretants]. According to Saussure’s first principle, the bond between the signifier [sign-vehicle] and the signified [interpretant] is arbitrary, or the signifier [sign-vehicle] is not motivated by its signified [interpretant]. It is thus plausible to believe that Bertin avoided the term “sign” in order to avoid arbitrariness, or as he termed it: conventionality. According to Bertin, ‘graphics has absolute natural laws. It is not “conventional”’ (Bertin, 1981: 177). Following Bertin, it is not arbitrary how the relation between sign-vehicles denote relationships: in order to denote a relationship of resemblance the sign-vehicles must transcribe a relationship of resemblance, in order to denote a relationship of order the sign-vehicles must transcribe a relationship of order; and to denote a relationship of proportion the sign-vehicles must transcribe a relationship of proportion (see Figure 3.4). Just as moving a chess-piece is dependent on certain rules, so is the transcription of the three relationships. To represent an a priori relationship of order, ‘it is necessary to use either a *value* variation from light to dark, or a *size* variation from small to large’ (Bertin, 1981: 145 – originally italics). In Figure 3.5 it is shown how the a priori relationship of order is transcribed by value variation.

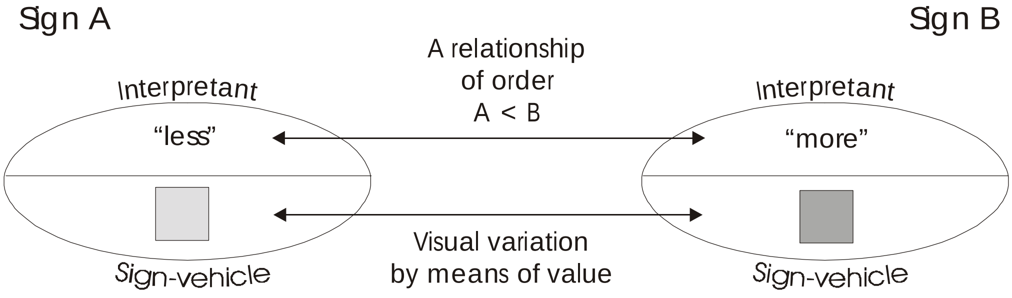


Figure 3.5: A dyadic sign model showing how the transcription of a relationship does not utilize signs [sign-vehicles] only the relationship between signs [sign-vehicles].

When elaborating how graphics transcribe a priori relationships, Bertin uses an example of representing a factory. 'How do we represent a factory?' Bertin asks. 'There is an infinitive number of "good" representations. The choice is an art. That is pictography' (Bertin, 1981: 177). Similarly, we could ask: How do we represent a municipality having a low level of employment in the industrial sector? There are several "good" representations, for instance a light shade of grey, red or blue. However, when representing relationships already defined, we must deal with graphics. 'Factory A employs twice as many workers as factory B. There is only one single representation: show that A is twice as large as B. This is not an art since *there is no choice*. This is graphics' (Bertin, 1981: 178 – originally italics). Similarly, municipality B has more workers in the industrial sector than municipality A. Following Bertin, to transcribe an a priori order between given sets (i.e. municipalities) 'it is necessary to use either a *value* variation from light to dark, or a *size* variation from small to large' (Bertin, 1981: 145 – originally italics). As the visual variable "value" is used for representing a low percentage of employment in industrial sector in municipality A, a darker "value" must be used to represent the higher percentage of employment in the industrial sector in municipality B. Consequently for graphics, the bond between the signifier [sign-vehicle] and signified [interpretant] is not arbitrary. There is indeed a reason for preferring a light shade to a dark shade when low values are to be represented and vice versa when high values are to be represented. Regarding contrary practices, Bertin claims: 'To represent the inherent *order* of quantities by a visual *non-order* or *disorder* of the signs is obviously a mistake and therefore gives a false image – in other words, false information' (Bertin, 1983b: 70 – originally italics). The bond between the signifier [sign-vehicle] and signified [interpretant] is in this respect fixed as their relationship must transcribe the a priori relationship to be understood as a graphic language. There is, however, no reason for preferring a light shade of red to a light shade of blue when representing low employment percentage in the industrial sector in municipality A. This might be regarded as conventions (i.e. arbitrariness) between certain mapping packages offering different colours for the symbolisation ramp (from light to dark within one colour).

The other position regarding the principle of arbitrariness (see Figure 3.2) is as a relation between the signifier and extra linguistic reality (reference). ‘Some of Saussure’s own examples as well as the history of the principle of arbitrariness suggest that arbitrariness is a matter of reference (Nöth, 1990: 243). One of the examples, which according to Helbig (1974: 40 – from Nöth, 1990: 61) expands the Saussurean diadic sign model to a triadic sign model because it includes a third element, is Saussure’s arguments that ‘the signified “ox” has as its signifier *b-ö-f* on one side of the border and *o-k-s (Ochs)* on the other’ (Saussure, 1974: 68). This third term, not included in the initial definition on the sign model, is according to Benveniste (1971: 44), the thing itself, the reality.

> It is only if one thinks of the animal *ox* in its concrete and “substantial” particularity, that one is justified in considering “arbitrary” the relationship between *böf* on the one hand and *oks* on the other to the same reality (Benveniste, 1971: 44, cited from Nöth, 1990: 244 – originally italics).

Thus, ‘Saussure’s argument of the arbitrary nature of signs necessarily requires reference to characteristics of objects in the world’ (Nöth, 1990: 61).

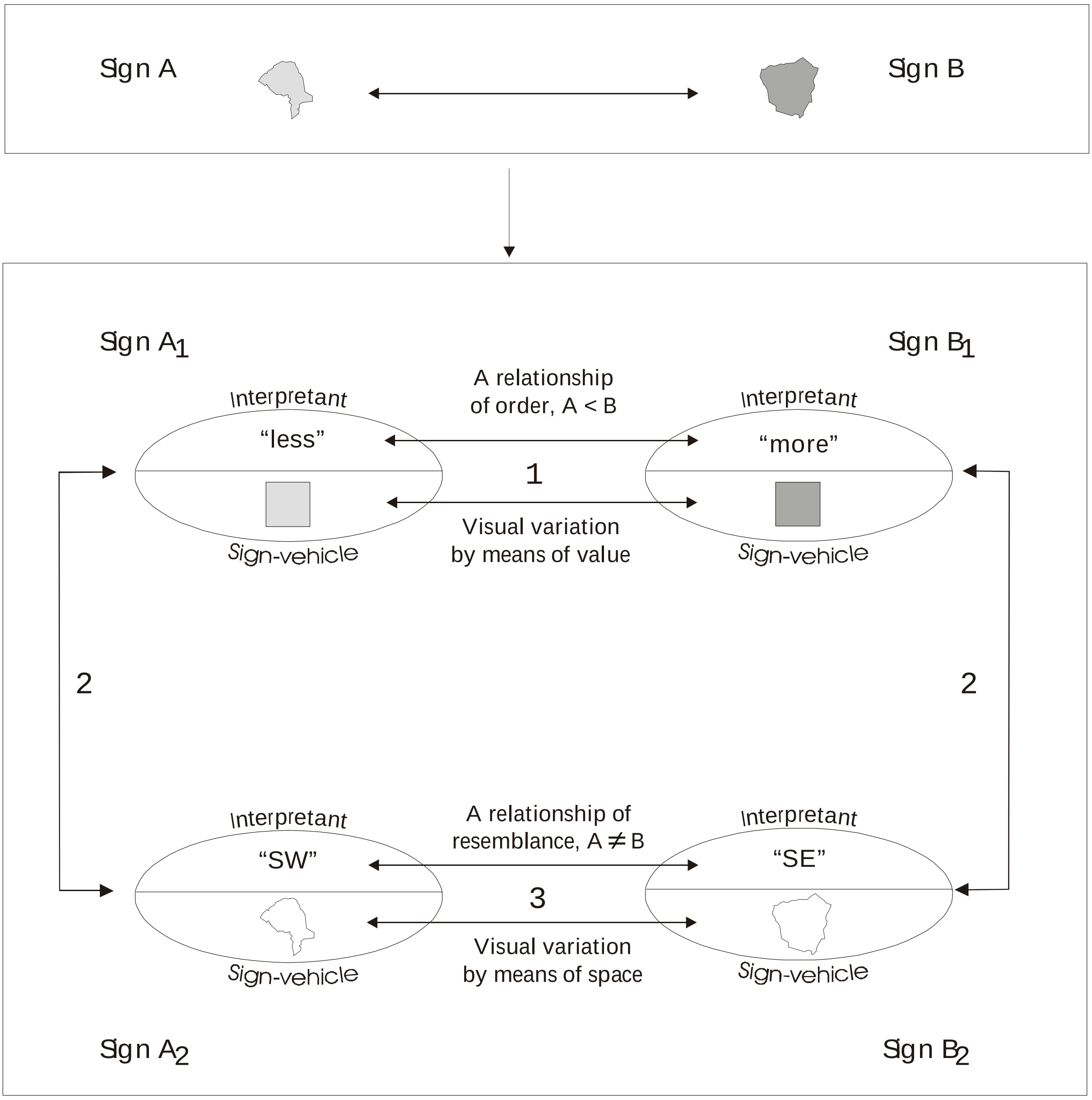


Figure 3.6: Decomposing the cartographic sign into its graphic sign and localised sign, which result in three types of syntactics for statistical maps: (1) denotative code, (2) interlocked code and (3) superimposed code (derived from Rød, 2001a, Figure 8).

A similar argument could be forwarded regarding cartographic sign-vehicles used for statistical mapping. However, it is necessary first to decompose the cartographic sign used in statistical mapping in its two components: the graphic sign, i.e. the retinal variable and the 'localised sign', i.e., the boundaries for a municipality (Rød, 2001a, Schlichtmann, 1990: 265). Two municipalities here referred to as SW (south west) and SE (south east), which are coloured with a light and dark shade of grey respectively constitute *three* types of sign-relationships as illustrated in Figure 3.6.

The three sign relationships depicted in Figure 3.6 is derived from Rød (2001a) where they are called (1) denotative code, (2) interlocked code and (3) superimposed code. In Rød (2001a), only the denotative and superimposed codes were discussed. The third sign-relationship, the interlocked code will be treated briefly below. All three sign-relationships are part of a map sign syntactics. Similar to Benveniste's argument, forwarding a cartographic argument about arbitrariness which includes a third element could read as follows: The signified [interpretant] "less" has as its signifier *light red* in one particular mapping package (e.g., ArcView) and *light grey* in another (e.g., GIB). It is only if one thinks of the state "less" in its descriptive particularity and its syntactical relation to "SW" (constituting its concrete particularity), that one is justified in considering "arbitrary" the relationship between *light red* on the one hand and *light grey* on the other to the same reality.

Developing this argument about the expansion of the diadic sign model to a triadic one may provide a more suitable sign model for interpreting the semiology of graphics. While Saussure claimed that there 'are no pre-existing ideas, and nothing is distinct before the appearance of language' (Saussure, 1974: 112) where language is to be understood as a system of values, Bertin claimed on the contrary that graphics, understood as a system of relationships, always start with a data table in where a priori relationships are defined. It is obvious that Bertin regards graphics as having reference to characteristics of objects of the world from how he defined graphics: 'Graphics involves utilizing the properties of the plane to make relationships of resemblance, order or proportions among **given sets** appear' (Bertin, 1981: 176 – my emphasis), and how he motivated graphics: 'The aim of graphics is to *make relationships* among **previously defined sets** appear' (Bertin, 1981: 176 – originally italics, my emphasis). It is thus plausible to regard the data table as the third element, the referent, existing before graphics, which expands the dyadic sign model illustrated in Figure 3.5 to a triadic one as illustrated in Figure 3.7.

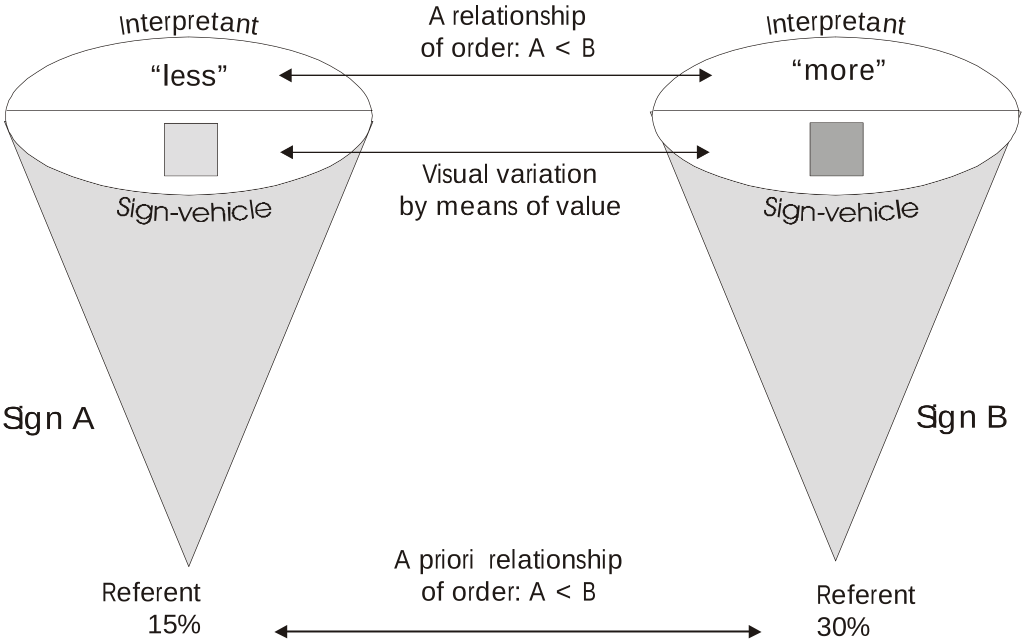


Figure 3.7: A triadic sign model showing how the transcription of relationships does not utilize signs; it utilises only the relationship between signs.

Saussure specified that the term arbitrary 'should not imply that the choice of the signifier is left entirely to the speaker' and he elaborates in brackets: 'the individual does not have the power to change a sign in any way once it has become established in the linguistic community' (Saussure, 1974: 68-69). The signifier 'is unmotivated i.e. arbitrary in that it actually has no natural connection with the signified' (Saussure, 1974: 69). However, individuals can indeed alter signs used for choropleth mapping, which suggests there is another sort of arbitrariness present: the assignment of sign-vehicles to referents. Depending on the degree of generalisation (i.e. the selected number of classes) and how it is performed (i.e. the selection of class borders), two different referents might have different or similar sign-vehicles. In Figure 3.7 the values 15% and 30% constitute the referents for sign A and sign B respectively, while the relationship between them constitutes an a priori relationship of order. 'On the choropleth map, data' e.g., the values 15% and 30%, 'will typically be classified resulting in relatively similar referents being assigned the same sign-vehicle, thus implying that they do not differ at all' (MacEachren, 1995: 266). It is arbitrary whether the a priori relationships are transcribed as a relationship of order (A < B as in Figure 3.7) or as a relationship of equal resemblance (A = B). The latter will be the case for a classification grouping the two values into the same class.

3.5.2 Direct and indirect signs

MacEachren characterises choropleth maps as indirect signs because they have indirect reference (MacEachren, 1995). I adopt the concept indirect sign for choropleth maps, but do not consider what Jenks (1967) proposed to call the data model to be the referent for the choropleth map as MacEachren does. I consider the data model and also the unclassed choropleth map as examples of direct signs since they both signify directly the values in the data table (see Figure 3.8).

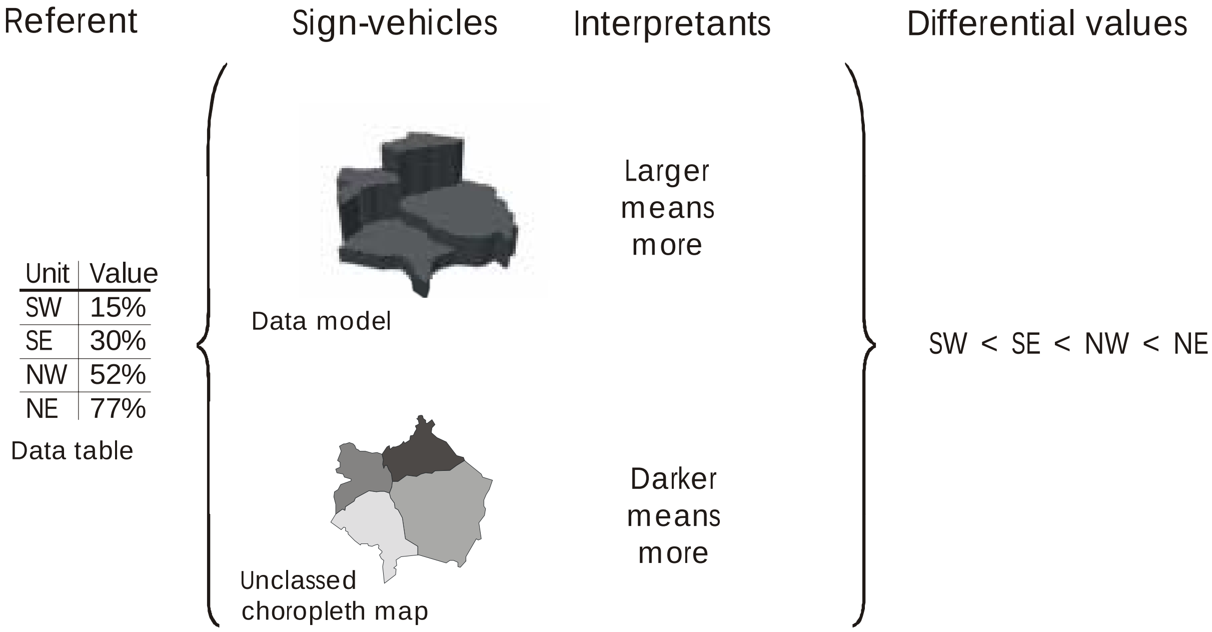


Figure 3.8: Direct signs have the same differential values.

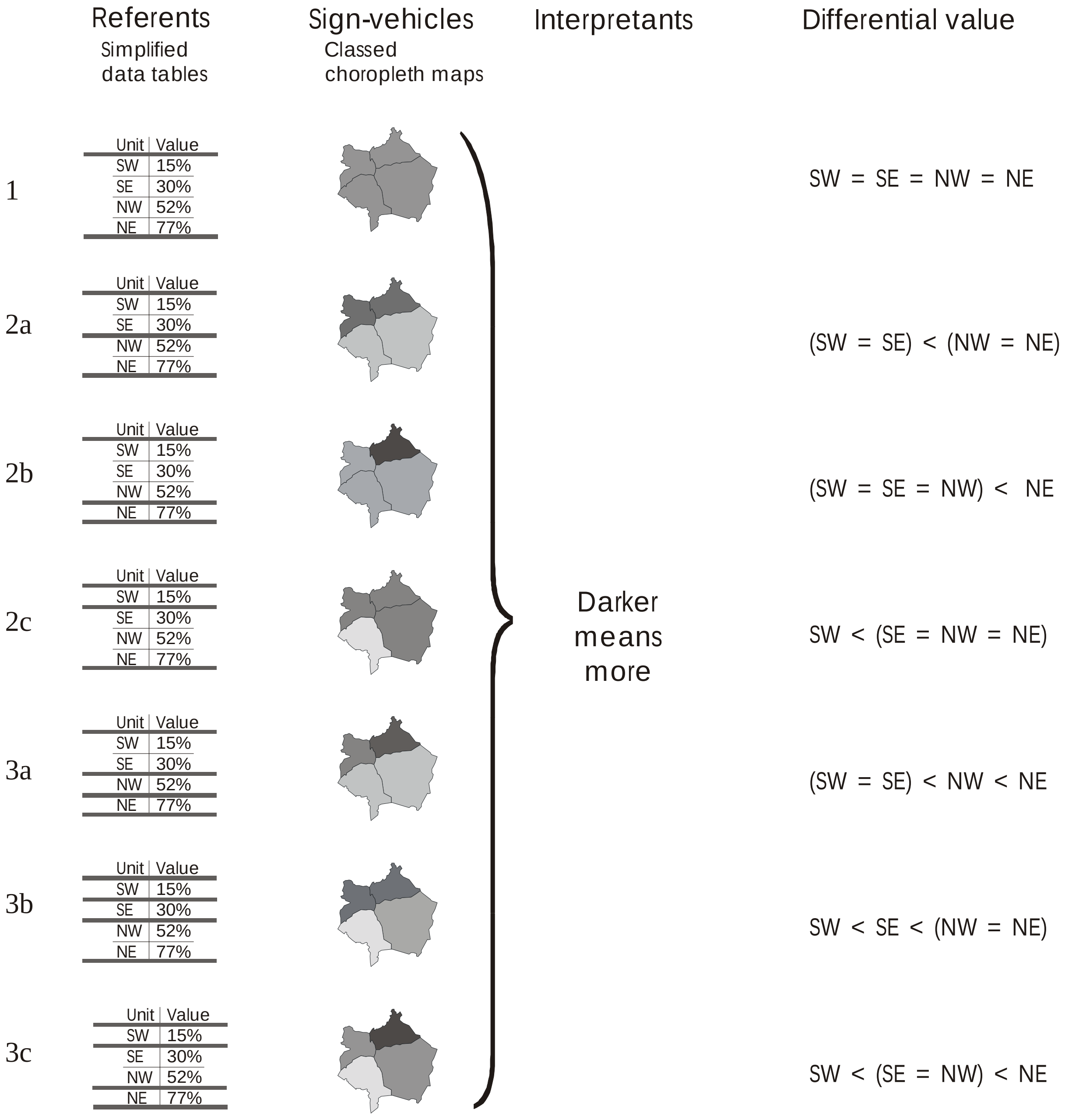


Figure 3.9: Indirect signs have various differential values.

I characterise these representational forms direct signs as their sign-vehicles signify the referent (i.e. the data table) directly and not via a simplified referent (i.e. a data classification). The data model and the unclassed choropleth map, however, do use different sign-vehicles and have different interpretants. Using Bertin’s term, the data model uses size variation while the unclassed choropleth map uses value variation. The data model has as its interpreter: “larger means more” and the *unclassed* choropleth map: “darker means more”. These two direct signs have the same differential values. The *classed* choropleth map is, however, an indirect sign as its sign-vehicle represents a *simplified* data table (see Figure 3.9). In Figure

3.9, the original referent is the same, but the various possible sign-vehicles refer to a classified data table in which the thick grey lines indicate class limits. All signs in this example have the same interpretants, “darker means more”, but they have various differential values.

While both direct signs in Figure 3.8 signify a relationship (differential value) which is identical with the a priori relationship: SW < SE < NW < NE, there are for the indirect signs in Figure 3.9 seven possible sign-vehicles showing various relationships (differential values) among the units, none of them identical to the a priori one:

- SW = SE = NW = NE (1)
- (SW = SE) < (NW = NE) (2a)
- (SW = SE = NW) < NE (2b)
- SW < (SE = NW = NE) (2c)
- (SW = SE) < NW < NE (3a)
- SW < SE < (NW = NE) (3b)
- SW < (SE = NW) < NE (3c)

Even if the number of classes is predefined, there is arbitrariness in the way the data is classified and thus in the way spatial patterns are generated in the maps and, consequently, in the expressed relationships among geographical units.

Making statistical mapping transparent means reducing this arbitrariness by enforcing the referential element in statistical mapping. For single choropleth map production, enforcing the referential element might be done either by advocating the use of unclassed choropleth maps or by using the classification method, which fits the data best for a given number of classes. Before outlining the methodology used in this thesis for making the statistical mapping process more transparent, some ethical implications of this arbitrariness will be considered.

3.6 Cartography, the real, and the good

The title of this section is borrowed from Sack (1999: 27) who gives an overview on the complex relation between ‘Geography, the real, and the good’. The relation between Cartography, the real, and the good, is here expressed very simply: ‘Cartography is about representation’ (MacEachren, 1995:1); it aims to represent the real, which in turn, might contribute to the good. Among Sack’s central perspectives on what is good, is ‘the quality of *seeing through to the real*' (1999: 34 – orginally italics), 'the value of a heightened and expanded awareness of reality' (1999: 34), and 'that everyone be given the greatest opportunity possible to know and expand his or her intellectual horizons' (1999: 39). 'As geographic beings, we are curious about the world and want to know what lies beyond the horizon' (Sack, 1999: 34). Russell points out that all knowledge is acquired either directly by perception of the external world, or indirectly by testimony (Russell, 1993: 103). 'The major part of our geographical knowledge cannot be verified by direct comparison with directly observed reality' (Keates, 1996: 139). If we care about expanding our intellectual horizon beyond its geographic-astronomic meaning as the limit where the sky meets the earth, we therefore do depend upon representations[9]. Maps generated by cartographic packages and GIS are representations and they may function as a way of constructing knowledge of the real.

Applying this perspective to cartography – if maps are to contribute to the good – they have to help people 'to see the world and its parts or places as clearly as possible and understand how these places make up the world' (Sack, 1999: 34). Several critical comments about how maps show the world and its parts of places have been forwarded refuting the idea that maps may contribute to the good. 'If we care about raising consciousness – then we need to face up to the conclusion that maps are often inadequate as a way of seeing' (Harley, 1990: 6). Inadequate because 'cartography communicates messages with a given objective, and colours these messages in order to meet that objective' (Ormeling, 1992: 62). Such mapping practices have a stamp of ethics as the maps tries to universalise a particular view of the world. 'When we make a map it is not only a metonymic substitution but also an ethical statement about the world' (Harley, 1990: 6). I believe, however, that to deliberately design maps to meet a certain objective is a rare activity. More common is the situation where the mapping packages do not allow for alternative cartographic designs but the 'one-map-solution' offered. As a result, instead of expanding an awareness of the real, cartographic representations may imbue a biased view of the world. For this reason, a disclosure of alternative cartographic designs is recognised as essential. I support Monmonier's critic of the 'one-map-solution':

> One-map solutions foster a highly selective, authored view perhaps reflecting consciously manipulative or ill-conceived design decisions about map scale, geographical scope, feature content, map title, classification of data, and the crispness or fuzziness of symbols representing uncertain features. But even if we are conscientious, even if we know our data inside and out, and even if we both know the creed of Bertin and Robinson and are aware of our own biases, the decision to present a **single** cartographic viewpoint can be a decision fraught with important ethical overtones (Monmonier, 1991: 3, originally emphasis).

When cartographic packages offer only one or a restricted number of possible maps, which is the case for the mapping module in Excel (see Rød, VI), it is an ethical failure. Among the six strategies forwarded by Monmonier for moving beyond the one-map solution, three of them are of special interest when implementing GIB. The first strategy is to allow the users to dynamically sequence through different cartographic views. The second complementary strategy is experimental mapping, allowing 'readers, users, or viewers to explore the data freely' (Monmonier, 1991: 4). Third, GIB aims to promote informed scepticism among students using it in geography and cartography courses by inviting them to see in how many ways the same data might be portrayed (i.e. mapped). Such scepticism is more likely to evolve if the human computer interface for cartographic visualisation is flexible and interactive.

As outlined in the previous sections, recognizing that maps have text-like qualities may elucidate how maps are imbued with meaning. The more practical answers to fundamental issues regarding cartographic representation or design will be treated in the next chapter on developing an interactive mapping package.

NOTES

[1] Harley did so by refering to McKenzie (1986: 35).

[2] This dichotomy is adopted by MacEachren (1995), but is also found in Bertin (1977/1981) who separated *graphic communication* from *graphic information processing* (Rød, 1998, Table 3).

[3] Several have supported Harley on this point, among them Miller (1992: 585) and Rundstrom (1993: 22), while other experiences of cartographers are as professionals with 'a subtle and critical sense of the nature of their work and not perceive cartography as an objective form of knowledge' (Godlewska, 1989: 97). I am in line with the latter opinion.

[4] The French word *mouton* translated to English can take two meanings: *sheep* or *mutton.* The value of the English term, thus, is different from the French one because 'English opposes *sheep* to *mutton,* while French does not have this difference in semantic value (Nöth, 1990: 61 – originally italics).

[5] In *How the monosemic graphics go polysemic* (Rød, 2001a) this premise is critically revisited regarding maps, resulting in the conclusion that diagrams and networks might be monosemic representations while statistical maps cannot.

[6] Bertin uses the notion "variables rétiniennes" or "retinal variables" for the six ways of variation a visual mark could take; variation in shape, orientation, colour, texture, value and size. Most often these six ways of variation are denoted in the Anglo-American literature as the visual variables when refering to Bertin. Bertin himself, however, used visual variables for both the retinal variables and the two dimensions of the plan and consequently, there are, according to Bertin, eight visual variables.

---

[7] Probably Nöth is here, as also Saussure sometimes did (Nöth, 1990: 60), using the term *sign* when referring to the signifier [sign-vehicle].

[8] Phrases in double quotation marks are citations from Saussure.

[9] Some might even claim that perception is a lower level of representation, (e.g.,visual perception is the representation of visual scenes before our eyes) and that cognition involves higher levels of representation (MacEachren, 1995:14).

4 Methodological approach

4.1 Introduction

Cartographic visualisation tools aid exploration, but they are designed for, and used exclusively by, experts. With this thesis, I want to forward arguments for a democratised cartographic visualisation that also allows exploration for ordinary users (Rød et al, 2001). The consequences of empowering non-specialists with statistical mapping opportunities are at once encouraging and alarming. Mostly, cartographers have asserted views that if everyone becomes a mapmaker this will result in poorly designed maps (Müller and Wang, 1990: 24; Weibel and Buttenfield, 1992: 223; Forrest, 1993: 144 and Kennedy, 1994: 16). My concern, however, is that poorly designed maps may not result, as most suggest, from new users' lack of familiarity with the principles of cartographic design, but because the software packages available tend to offer rather arbitrary 'one-map-solutions' (Monmonier, 1991). "One"- or "few-map-solutions" are offered by democratised mapping packages, e.g. the mapping module in MS Excel, because of a prerequisite for simplicity software usage, which is obtained by stripping flexibility and interactivity (Rød, 1999). Flexibility is here understood as the ability to choose among several functions, e.g. methods for data classification and freedom in adjusting these functions, e.g. by manipulating class breaks. Interactivity is understood to be a close interplay between user and software package where the program during processing presents results and can respond to new instructions or load new data.

Making a choropleth map in the MS Excel mapping module is indeed simple. After having selected the cell references, stored in a spreadsheet containing the data to be mapped, a choropleth map according to a quintile classification appears as default representation. Consequently, users do not need to worry about the essential decisions underpinning the design of a choropleth map and should as such either have a map, which *presents* their data ready made or be able to devote more time to *analysing* their data by, for instance, studying the map. Unfortunately, this is not the outcome of the simplified use of this mapping package. In general, a high degree of user interaction is recognised as paramount for visualisation tools to succeed (MacEachren and Ganter, 1990: 74), and, I would add, flexibility is very important for communication devices to succeed. Users of visualisation tools ought to be able to discover yet unknown geographic knowledge while users of communication devices ought to be able to adjust the design according to the message the map is supposed to transfer. While the MS Excel mapping module might be regarded as interactive regarding the few functions included, the software is too constrained to function as a visualisation tool. In terms of class interval selection, using the other classification methods offered, i.e. equal intervals, is the only way to alter class intervals. One is constrained from manipulating class breaks, which might be necessary to investigate areas above or below certain values. Consequently, it is unlikely that the choropleth map resulting from using Excel would be of much use either for presenting or for analysing data.

While developing GIB, it has become evident that I commit myself both to the communication paradigm and the visualisation perspective in cartography. One of the central aims for the communication paradigm was to find “the optimal map” and alongside GIB, I am developing it to “optimise” data classification in order to produce “an optimised choropleth map”. Still, as other devotees to the visualisation paradigm with similar approaches (Egbert and Slocum, 1992; Andrienko and Andrienko, 1999), I am developing a mapping package highly interactive regarding class break selection. The users of GIB might alter class breaks for choropleth maps in numerous ways and will always have responses to their choices. Each of the constructed choropleth maps are provided with responses on how well the particular data classification “fits” the original data distribution, i.e. how significant differences between observations might disappear when the observations fall within the same class and/or how insignificant differences between observations might appear emphasised when the observations fall into adjacent classes (see Figure 4.1).

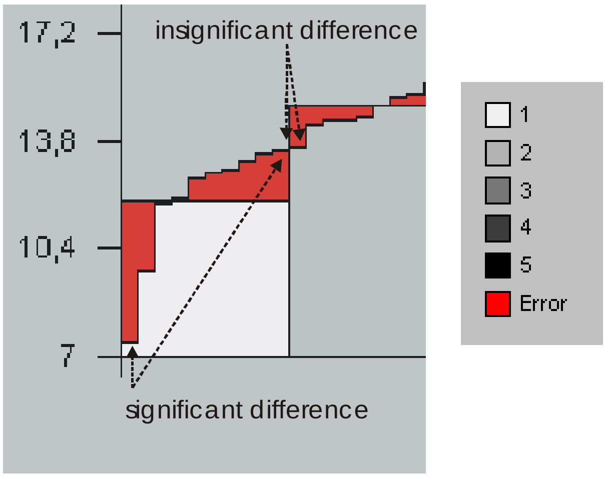


Figure 4.1: Class break resulting in suppressing a significant difference and emphasising an insignificant one (from the graphic array generated by GIB – see manual section 4.3).

If users do not need to visualise their data idiographically, e.g. find out how the observations are grouped according to particular values such as mean and standard deviations, they are recommended to classify their data using the optimal method, simply because the more the choropleth map is accurate the less the statistical phenomena is misrepresented. Classification accuracy is here understood as how well the information contained in a map corresponds with the information provided by a data table – the better the correspondence, the better, or more accurate, is the classification.

In this superstructure, I will not differentiate (as I did in Rød, 1998) between objective and subjective criteria for good classification and devote these criteria to the communication paradigm and the visualisation perspective respectively. Instead, I will promote a combined criterion for good classification. Indeed, it is important for the users to be able to put their data into different perspectives, but the statement that a 'criterion for good classification is whether the map shows the geographical pattern the user is looking for' (Rød, 1998: 41) – I will totally reject. Krygier gives a good reason why. While working on his Ph.D. thesis, 'Krygier recounts an experience in which a client (a historian) objected to "optimal" data classification because it did not reveal the pattern that she "knew" was there' (MacEachren 1995: 210). The "optimal" map is the map which most accurately portrays the original data according to some statistical criterion and as stated in Rød (2000): if the classification is accurate it is more likely that the pattern created refers to a real world situation. Thus, an accurately classified map decreases the probability that geographical differences, which are present in the data set, will be suppressed in the map and/or that geographical differences hardly existing in the data set will be created or exaggerated in the map.

Two central questions have arisen while implementing GIB.

1. Is it necessary, as in the dominant user interfaces for democratised GIS, to make the process of statistical mapping simple by stripping functionality; or might simplicity be obtained by other means?
2. Is it possible to initiate a critical reflection on the way a choropleth map represents a statistical phenomenon?

I share the views expressed by Egbert and Slocum that software packages for statistical mapping seem to assume that the resulting map will be used like a traditional paper one

(1992: 275). This seems evident because these program packages do not offer flexible and interactive graphics enabling users to adjust the map design to better communicate a certain message or to explore the database underlying a map. GIB is developed in this vein permitting users to explore the data freely through multiple cartographic representations. Still, simplicity is maintained by offering users default options based on rules as well as an interactive guidance by on-screen choices, resources to make the choices, and appropriate program responses.

While an approach to offering only a limited number of choropleth map solutions may well be called autocratic, an approach on offering a multitude of choropleth map solutions might result in relativism. Users of GIB, who are offered a multitude of choropleth map solutions, might experience, if they interactively try out different class intervals, that the resulting maps will express divergent political standpoints (like in Figure 4.5), and consequently, realise that ‘all maps state an argument about the world’ (Harley, 1989: 11). In order to avoid relativism and to offer some grounds from which to choose among “competing truth claims”, the system performs calculation on classification accuracy (see Rød VI), which is used as a rule base when offering default options on, for example, classification methods for choropleth maps. The development of GIB has been an attempt to ‘merge expert system and human interaction to facilitate knowledge discovery’ (MacEachren, 1995: 433). It is also hoped that users will discover that for numerous possible cartographic representations of one statistical variable, their maps possess certain accuracy in the manner they hide significant variation or exaggerate insignificant variation (see Figure 4.1). The same calculations on classification accuracy are thus applied to make the statistical mapping process more transparent.

4.2 Developing rule bases for statistical mapping

The fear of the negative consequences of a democratised mapping has fed many arguments to the need of developing expert systems in order to assure the quality of the resulting map. If users do not have knowledge about good map design, it ought to be imbedded in the software so that the system may take the essential decisions (Weibel and Buttenfield, 1992: 224). Expert systems are computer programs constructed to emulate decisions an expert would have taken according to a certain amount of information about a problem. Ideally, the decisions should resemble the one taken by an expert. Kraak and Ormeling (1996: 211) give some general guidelines for the implementation of a rule-based expert GIS which can be summarised accordingly:

1. To offer the users logical and reasonable default options. Often the users of the GIS or cartographic software are not interested in design rules and will regard it as convenient to follow the default offered. Therefore these default options should be based on cartographic rules.
2. To offer the users wizard-functions, accompanying every stage of the map design process by a set of suggestions. The use of a wizard-function will make the mapping process flexible since the user might move back and forth and redo her choices.
3. In particular situations, warnings or other suggestions should be forced upon the user when the user is about to violate cartographic rules.

Certain aspects from the above are implemented in GIB.

1. As suggested by Rød et al (2001), a rule base founded on the level of measurement for a certain variable has been developed, which constrains the type of map offered as default for this variable. More profoundly developed is a rule base founded on statistical accuracy, which activates when users select a method for data classification when preparing choropleth maps (Rød, VI). Classification accuracy is calculated for each method and the one producing the most accurate map becomes the default option (see Figure 4.3a). For arithmetic and geometric progressions, which require additional decisions regarding whether they should be concave or convex and whether they should be used with a constant, increasing or decreasing rate, the combination of options giving the best fit is offered as default (Rød, VI: 12).
2. GIB offers users producing a choropleth map a three stepped wizard function corresponding with the steps in the choropleth mapping process (Baudouin, 1987: 323; Cromley, 1995: 15):
   i. The selection of the number of classes.
   ii. The selection of how breaks between classes are set.
   iii. The selection of a graphic symbolisation scheme for the number of classes.

   Suggestions for each of these stages are accompanied in the GIB mapping package. The users might move back and forth and redo their choices.
3. Users are able to produce four types of statistical maps using GIB: choropleth maps, dot density maps, proportional point symbol maps and chorochromatic maps. If users try to make a type of map, which does not correspond with the variable’s level of

measurement, they are prompted a warning reminding the users on the link between measurement level and appropriate cartographic representation.

4.2.1 Measurement level as rule base

Stevens defined measurement as the assignment of certain properties (numerals) to objects (attribute numbers) or events according to rules (Stevens, 1946: 677). This thesis follows Steven’s classification of measurement levels, i.e. nominal, ordinal, interval and ratio, but makes a distinction between *relative* ratio and *absolute* ratio in the same manner as Bertin (Bertin, 1981: 190). It is a complex problem to design rules that automatically assign the measurement characteristics matching a particular variable (Wang and Ormeling, 1996). In the GIB mapping package, after users have selected a variable to be mapped, they were asked to indicate its measurement level. From the five options: nominal, ordinal, interval, relative ratio and absolute ratio, a default option is provided based on these rules:

- if the values are characters, the measurement is set to nominal
- if the values are integer numbers between 0 and 10, the measurement level is set to ordinal
- if the values are decimal numbers the measurement is set to relative ratio.
- if the values are integer numbers exceeding the range between 0 and 10, the measurement is set to absolute ratio

The interval level is never offered as a default option.

This implementation was done in order to generate a reflection on the importance of knowing the level of measurement for the variable before making any choices on cartographic representation and cartographic symbolisation. These issues on using the measurement level as a rule base are elaborated further related to cartographic representation in Rød et al (2001) and to cartographic symbolisation in Rød (1998). The compulsion for the users of GIB to respond to a question on measurement level, which I hoped would inform them of the link between measurement level and appropriate cartographic representation, rather seemed to be another boring task users needed to react to. Consequently, this has been altered. To improve the understanding that certain cartographic representations fit data whose values are on a certain level of measurement, immediately after a variable is selected, GIB analyses the data values, indicates their level of measurement and generates an appropriate statistical map according to the following schema:

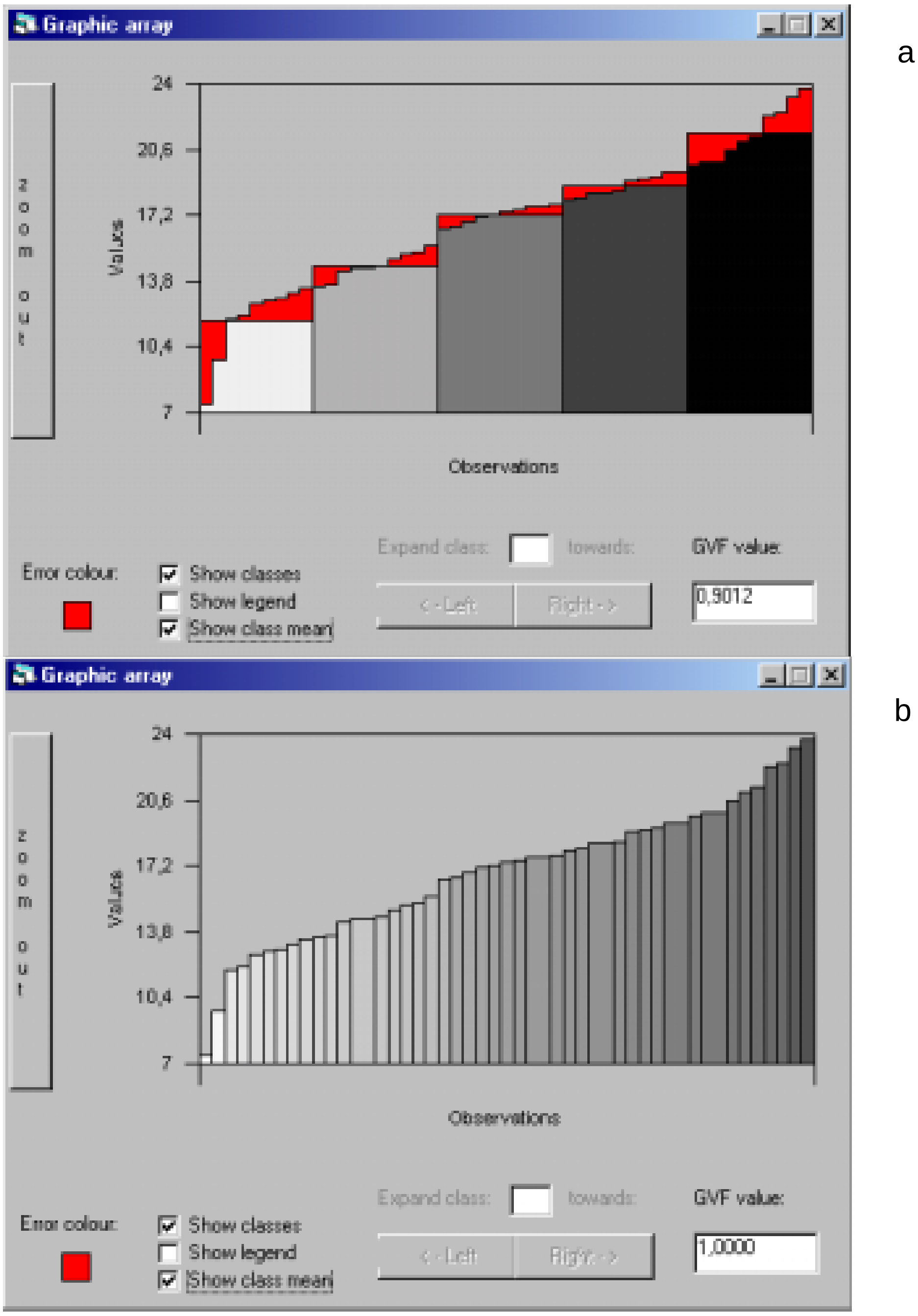
nominal level of measurement Ą chorochromatic map
ordinal level of measurement Ą unclassed choropleth map
relative ratio level of measurement Ą unclassed choropleth map
absolute ratio level of measurement Ą proportional point symbol map

The unclassed choropleth map is deliberately chosen as the default representation for ordinal, interval and relative ratio level of measurement. GIB is implemented to generate class less choropleth maps in accordance with the suggestions made by Kennedy (Kennedy, 1994: 24). Choropleth maps without class intervals ‘on which the visual intensity is exactly proportional to the data intensity’ have ‘no quantization error’ and consequently, the difficult problem of optimum class intervals are thus circumvented’ (Tobler, 1973: 262). In order to encounter the main argument against unclassed choropleth maps – degenerated readability – GIB displays the value of the geographical units which is brushed over by the cursor.

4.2.2 Data classification accuracy as rule base

The most accurate choropleth map that can be made from a certain data set is an unclassed choropleth map (i.e. a choropleth map where the number of classes equals the number of unique observations values). The most inaccurate choropleth map that can be made from a certain data set is a one-class map. Classification accuracy, thus, increases with an increasing number of classes, which is illustrated in Figure 4.2 as a matching problem between the original data values and the class means.

Figure 4.2 shows the original data values represented by bars sorted according to increasing values and shaded according to the class the observation belongs to. The sum of deviations between observation values and class means within each class is denoted as the classification error and depicted as an area coloured with a certain “error colour”. If the observation values within each class are similar, the deviations are small and consequently, only a modest classification error is present (e.g., class three in Figure 4.2a). However, if there is considerable within-class variation, the classification error becomes significant (e.g., first and fifth class in Figure 4.2a). A classification principle optimalising the choropleth map tries to increase the within-class homogeneity and the between-class heterogeneity. An index expressing the classification accuracy is generally known as the *goodness of variance fit* – GVF (see Rød, 1999; 2000; VI) and is used in the GIB mapping package. For unclassed choropleth maps where there is only one observation – or two or more observations with identical values – per class, the deviations between observation values and class means are zero and thus, there is no classification error and the GVF value equals one (see Figure 4.2b).



a

Error colour: 🟥

☑ Show classes
☐ Show legend
☑ Show class mean

GVF value: 0,9012

b

Error colour: 🟥

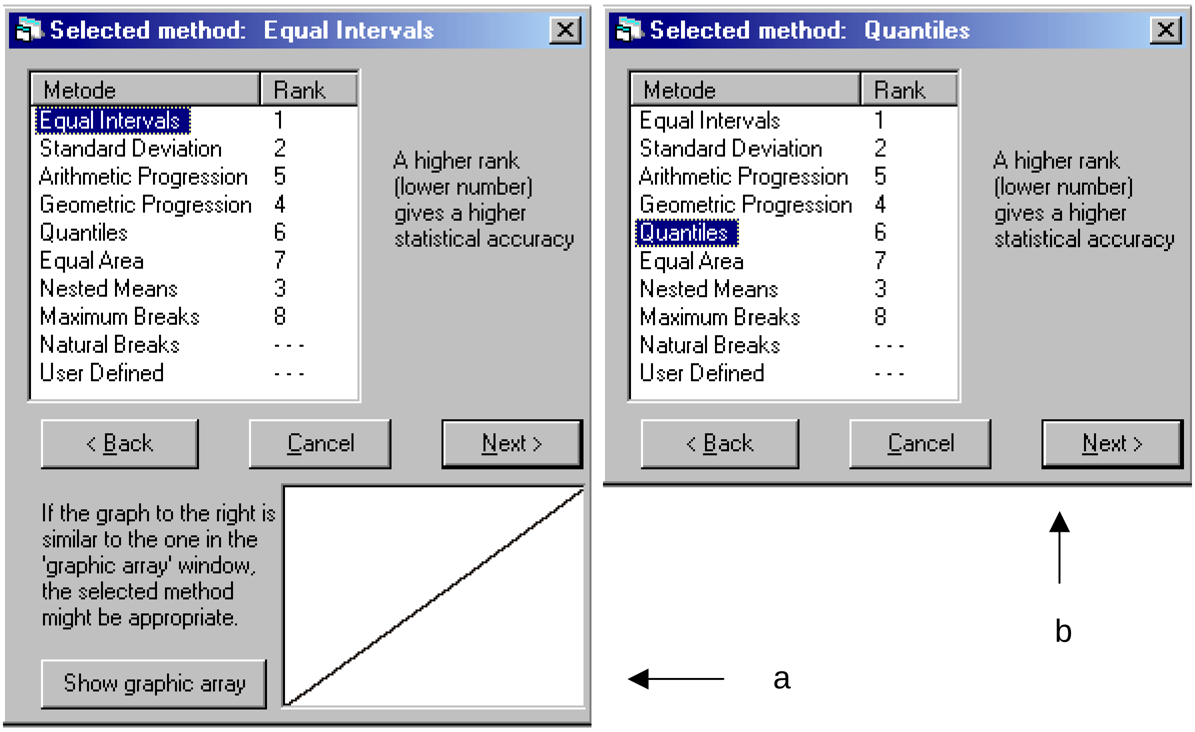
☑ Show classes
☐ Show legend
☑ Show class mean

GVF value: 1,0000

Figure 4.2: Sorted observation values for a certain variable (the percentage of population above 67 years in Trondheim in 1997). (a) Values are classified into quintiles (five near-equal classes) produce classification error. (b) Values are “unclassed” and, consequently, it is a perfect fit between the original data values and the “classification”.

In the GIB mapping package data classification is optimalised in two ways:

1. When selecting a method for data classification, the method producing the best "fit" is offered as default and the alternative methods are ranked according to their "fitness" (see Figure 4.3).



Selected method: Equal Intervals

Metode	Rank
Equal Intervals	1
Standard Deviation	2
Arithmetic Progression	5
Geometric Progression	4
Quantiles	6
Equal Area	7
Nested Means	3
Maximum Breaks	8
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

If the graph to the right is similar to the one in the 'graphic array' window, the selected method might be appropriate.

Selected method: Quantiles

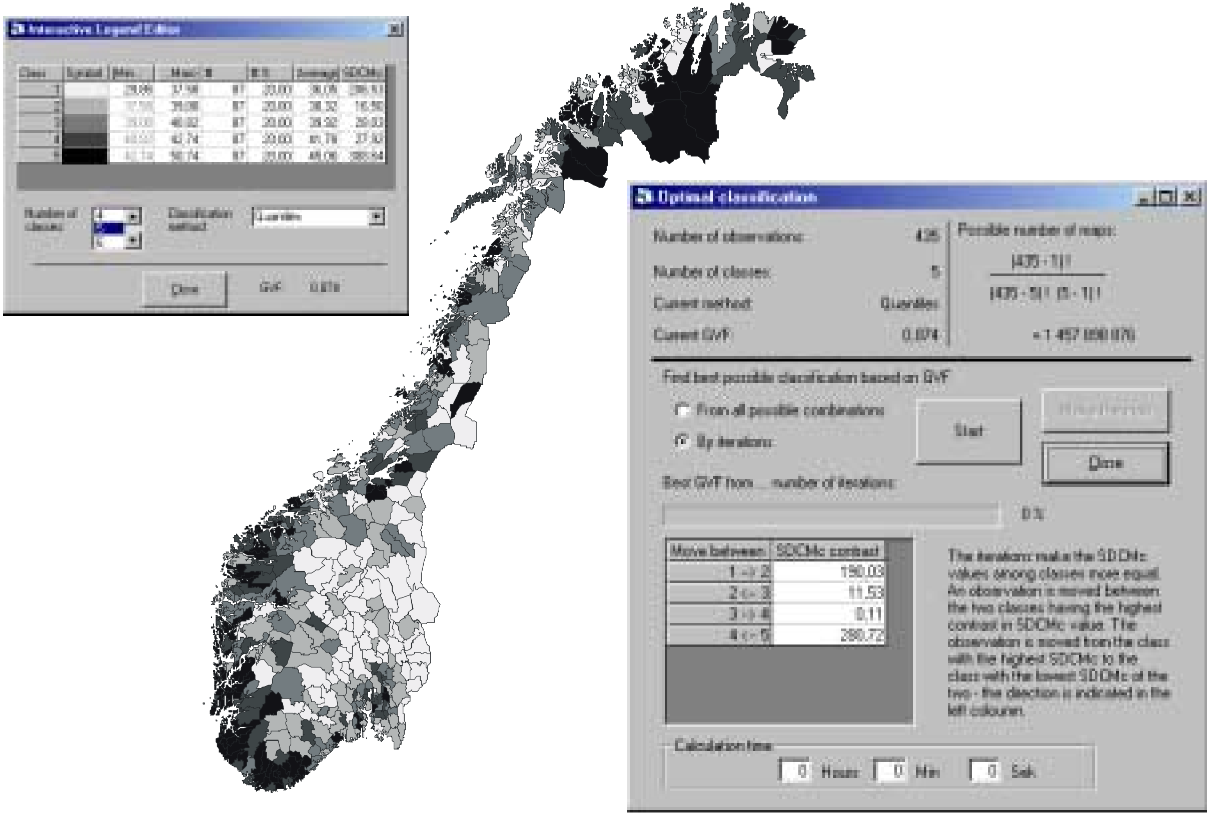
Metode	Rank
Equal Intervals	1
Standard Deviation	2
Arithmetic Progression	5
Geometric Progression	4
Quantiles	6
Equal Area	7
Nested Means	3
Maximum Breaks	8
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

Figure 4.3: Dialog box in GIB for selecting classification method. All methods except natural breaks and user defined are ranked according to their GVF scores. The method resulting in the highest GVF is ranked as number one and is offered as default. (a) For all the serial classification methods, an ideal graph for visual comparison is shown. (b) When idiographic methods are selected, the dialog box becomes reduced.

2. As an optimal classification method where the highest within-class variations (denoted as SDCMc) are tried to be reduced by moving observations from one class to another. An observation is moved between the classes where the contrast in SDCMc values is highest (see Figure 4.4). The observation is moved <u>from</u> the class having the highest SDCMc value to that which has the smallest SDCMc value. Iterations on this moving operation continue until the GVF value does not increase more than 0.001[1]. In Figure 4.4 it is shown a situation in where the GVF index does not increase more than 0.001 after 200 iterations, but then the SDCMc values are more balanced.

Before optimalisation



Optimal classification

Number of observations: 435
Number of classes: 5
Current method: Quantiles
Current GVF: 0.874

Possible number of maps:

$$\frac{(435-1)!}{(435-5)!\,(5-1)!} = 1\ 457\ 898\ 876$$

Find best possible classification based on GVF
○ From all possible combinations
● By iterations

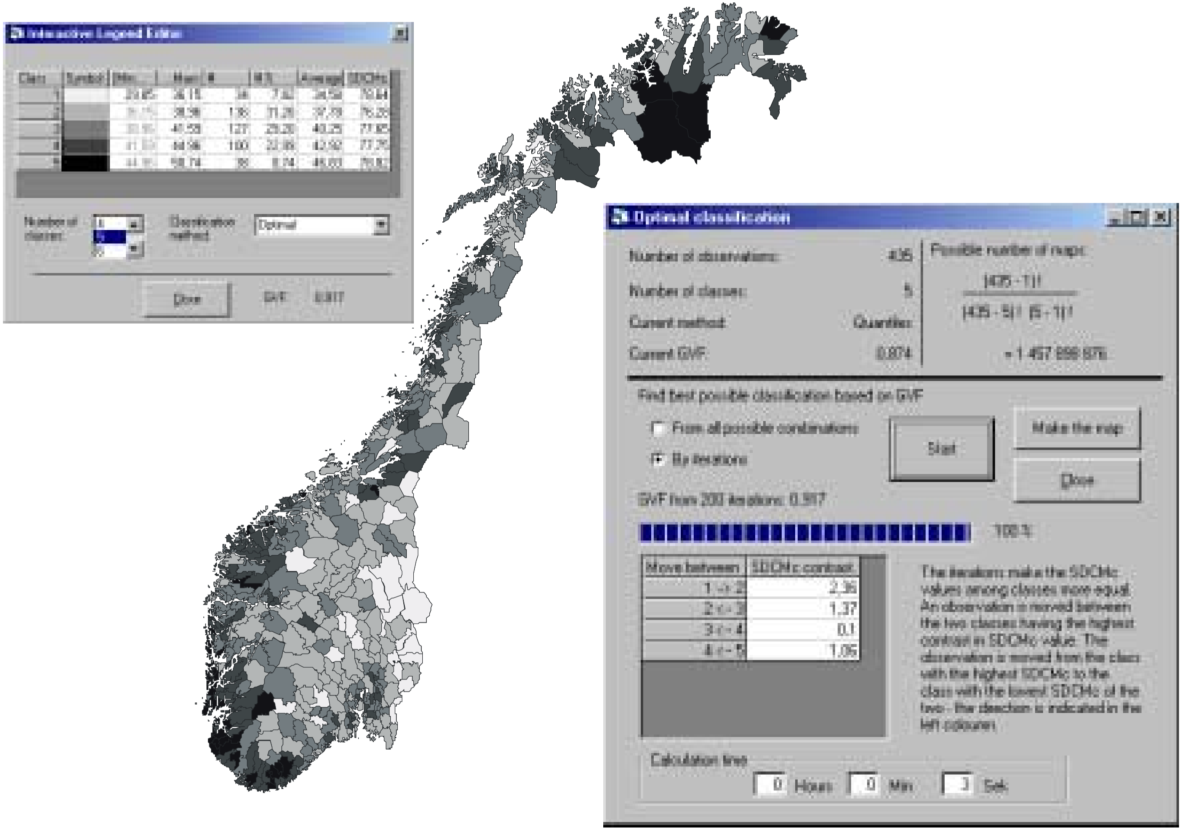
Best GVF from .... number of iterations:
0 %

Move between	SDCMc contrast
1 -> 2	190,03
2 <- 3	11,53
3 -> 4	0,11
4 <- 5	280,72

The iterations make the SDCMc values among classes more equal. An observation is moved between the two classes having the highest contrast in SDCMc value. The observation is moved from the class with the highest SDCMc to the class with the lowest SDCMc of the two - the direction is indicated in the left colounm.

Calculation time
0 Hours 0 Min 0 Sek

After optimalisation



Optimal classification

Number of observations: 435
Number of classes: 5
Current method: Quantiles
Current GVF: 0.874

Possible number of maps:

$$\frac{(435-1)!}{(435-5)!\,(5-1)!} = 1\ 457\ 898\ 876$$

Find best possible classification based on GVF
○ From all possible combinations
● By iterations

GVF from 200 iterations: 0.917
100 %

Move between	SDCMc contrast
1 -> 2	2,36
2 <- 3	1,37
3 <- 4	0,1
4 <- 5	1,06

The iterations make the SDCMc values among classes more equal. An observation is moved between the two classes having the highest contrast in SDCMc value. The observation is moved from the class with the highest SDCMc to the class with the lowest SDCMc of the two - the direction is indicated in the left colounm.

Calculation time
0 Hours 0 Min 3 Sek

Figure 4.4: Dialog boxes for optimal classification and interactive legend editor and the resulting maps before and after optimalisation. The maps show the percentage of population aged 0-20 years in 1997 (see Rød, 1999 and chapter 6).

4.3. Expert system versus experienced learning

While every stage in the design of a choropleth map used for presenting well-defined messages could probably be automated using an expert system approach, choropleth mapping used in an exploratory manner cannot. ‘Clearly only a small subset of maps are produced to “communicate” a particular message’ (MacEachren, 1995: 6). Only taking only an expert system approach is thus insufficient. Futhermore, an expert system approach is based upon a confined comprehension of ordinary people’s capabilities. As new map-authors may not have what cartographers might consider as the necessary cartographic knowledge, this knowledge ought to be embedded in the system. I find that such arguments too readily fail to take into considerations these users’ abilities to reason. In order for cartographic packages to contribute to an increased geographic understanding, they have to allow an iterative cognitive cycle of ‘seeing that’ and ‘reasoning why’ (MacEachren, 1995: 363). I believe, like MacEachren and Ganter, that ‘rather than developing expert systems that help find a single optimal map for representing a set of information, we need to develop systems that encourage exploration of multiple perspectives on the same data’ (MacEachren and Ganter, 1990: 75). In *An agenda for democratising cartographic visualisation* (Rød et al, 2001), we take an experienced learning approach. ‘The notions *expert* and *experience* both imply dealing with knowledge. Approaches to expert systems tend to concentrate on the necessity of embedding expert knowledge, while an experienced learning perspective concentrates on the possibilities for anybody to gain new knowledge’ (Rød et al, 2001: 38 – originally italics). This might be a naïve point of departure, but instead of viewing the extension of the group of mapmakers as threatening proper map making it might better elucidate another possible outcome of the democratisation of mapping. When the map becomes a general medium for exploring and expressing opinions, just like any other textual or oral expression, critical voices towards these representations will also evolve. This is recognised here as a beneficial outcome of democratisation cartographic visualisation.

4.4 A flexible and interactive cartography

The selection of class intervals can strongly affect the visual impression given by a map. The many different sets of interval breaks available for most classified maps illustrate how easy it is to distort the portrayal of the underlying numerical distribution, and, in the process, to propagate “lie” with maps.

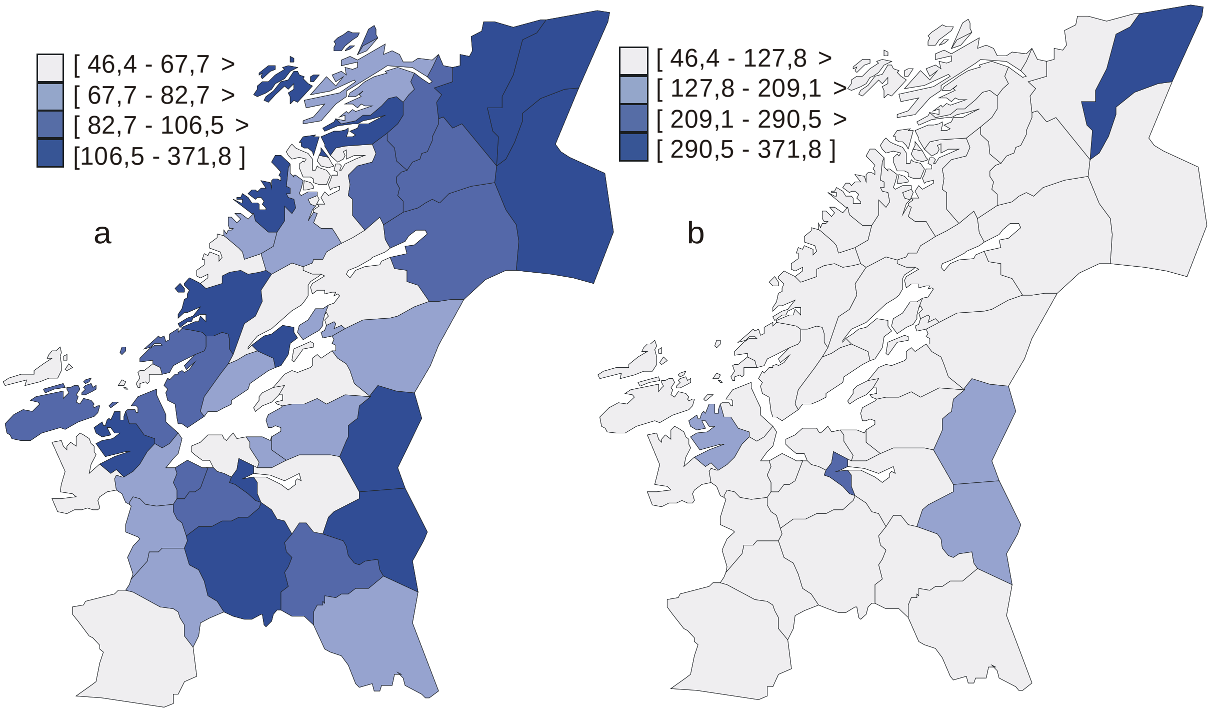


Figure 4.5: By selecting two different principles for class intervals: (a) quantiles and (b) equal intervals, two different portrayals of *the number of places in institutions for the elderly per 1000 inhabitants above 67 years in Trøndelag in 1997* result. (From Rød, VI, Figures 5 and 6). The map on the left emphasise a high welfare level while the map on the right emphasise a low welfare level.

Figure 4.5 is one example of how differently two maps can portray the same variable; here the number of places in institutions for the elderly per 1000 inhabitants above 67 years in Trøndelag in 1997. The two maps differ only in their classification method. The two most common classification methods are applied: quantiles and equal intervals respectively. Obviously, ‘class breaks can be manipulated to yield choropleth maps supporting politically divergent interpretations’ (Monmonier, 1991b: 41). If the intention is to promote high welfare for the elderly, Figure 4.5a should be selected and, conversely, if one rather would like to express a situation of low welfare for the elderly, Figure 4.5b should be selected.

GIB offers an interactive legend editor allowing users to change the number of classes, classification method and class breaks and thus they are able to produce a multitude of alternative visions (see Figure 4.1.2 in the GIB manual). Additionally, for each of these visions, users are informed about the degree in which the map “lies”; how information is hidden, and how information is visually exaggerated. This is done visually by coloring “error areas” in a graphic array[2] and numerically by showing index values in the interactive legend editor indicating the accuracy within classes (SDCMc) as well as the overall accuracy (GVF). A detailed outline on the visual and numerical responses on error in choropleth maps is provided in *De nauwkeurigheid van geclassificeerde karten* (Rød, 2000). The most accurate map is the unclassed choropleth map. The alternative is to offer a variety of classification options to improve the correspondence between the original data and the mapped representation (see Rød, VI). The two instruments that are developed in GIB that aid the decision of class intervals are outlined. For the serial types of classification methods (see Evans, 1977: 101-102 and Rød, VI: 9), users have the opportunity to visually compare the distribution of the variable with the ideal form it should take if a particular method is appropriate (see Rød, VI, Figure 2).

4.5 Prototyping a software package for statistical mapping

Prototyping is one of several methodologies for information systems development. Prototyping is sometimes referred to as *the iterative development approach* (Sølvberg and Kung, 1993).

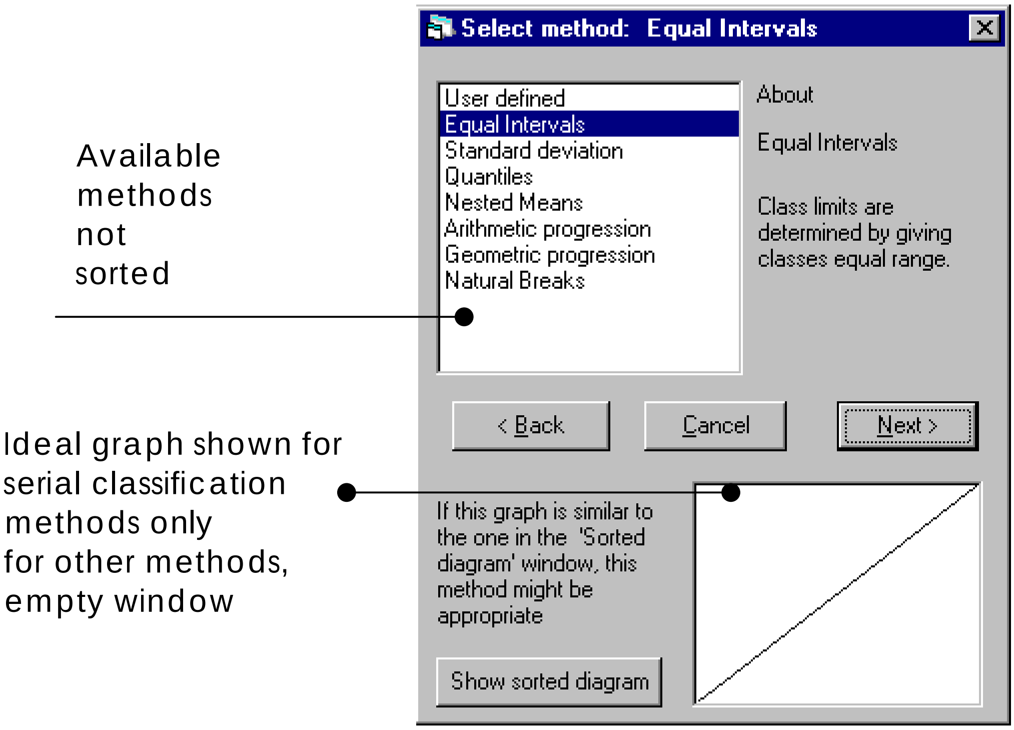
> The main idea is that the most important operational functions of an information system are designed, implemented, installed and put into operation as quickly as possible. The system’s evolution is seen as a sequence of addition and modifications to the specifications and to the software, as users’ operational experience are forcing system changes (Sølvberg and Kung, 1993: 425).

Prototyping essentially involves five phases leading to stepwise refinements until the prototype becomes the system (from Hirschheim et al, 1995: 242):

1. Identifying some of the basic requirements without any pretensions that these are either complete or not subject to drastic changes.
2. Develop a design that meets these requirements and implement it.
3. Have the user experiment with the prototype noting good and bad features.
4. Revise and enhance the prototype accordingly thereby redefining and gradually completing the requirements and also improving the interface and reliability.

Repeat steps 3 and 4 until the user is satisfied or time and money foreclose on further revisions.

As GIB has been used in a geography course for first year students, it has been possible to observe the prototype, noting both good and bad features. As part of the curriculum for the first year students studying geography, students must undertake a collaborative project on the socio-economic development of a Norwegian county. They thus need to produce time-series maps that may reveal geographical changes. The most appropriate types of map to use for this requirement are choropleth maps classified after some exogenous classification scheme making comparison between the maps possible, most often by applying 'a data-calibrated system to the *combined* data of the two or more periods of variables' (Evans, 1977: 108 – originally italics). Unfortunately, the individual maps from these time series are in no way optimised for the data portrayed: thus, the need for map production do not match the main requirements underpinning the development of GIB. Still, several good and bad features related to the more general operational functions of GIB were identifiable, but fewer related to the core features, like classification accuracy applied for user guidance and quality control on class intervals selection. Therefore, a small group of second year geography students was observed more closely according to the “think-aloud-method” (Someren et al, 1994) in order to identify the good and bad features related to GIB’s core functionalities. The “think-aloud-method” 'is a very direct method to gain insight in the knowledge and methods of human problem-solving' (Someren et al, 1994: 1). 'Using this method the subjects are asked to concurrently give a running commentary on their thoughts and decisions as they complete the task; this commentary is recorded verbatim and then analysed' (McGuinness, 1994: 187). The method was applied to observe user experiences with GIB to obtain guiding principles in the building of a rule-based system for statistical mapping. A small group of students was observed while they were making choropleth maps of variables with different distributional character. The most promising outcome of this test was the design of a dialog box where users select the method for data classification (see Figure 4.3 and 4.6). Although the students had little or no knowledge about particular classification methods, they were all able to select an appropriate method based on a simple visual inspection. For a highly skewed variable about places in institutions for the elderly, they all selected an arithmetic or geometric progression as classification method. However, a visual comparison makes sense only for the serial classification methods because these have an ideal form that the mapped variable’s distribution should resemble if the particular method is adequate. The students became confused as the ideal distributional forms were only shown for some methods – the serial ones – and not for others – the exogenous and idiographic ones (see Figure 4.6).



Select method: Equal Intervals

About

Equal Intervals

Class limits are determined by giving classes equal range.

If this graph is similar to the one in the 'Sorted diagram' window, this method might be appropriate

Available methods not sorted

Ideal graph shown for serial classification methods only for other methods, empty window

Figure 4.6: Dialog box evaluated by the “think-aloud” method. Users were able to select from a variety of methods, but as only the serial methods gave a graph in the lower window, this confused the users. As an additional guidance to the visual one, the GVF index were calculated for each method in order to rank the methods and to offer the best of them as default (see Figure 4.3).

Instead of leaving the window empty for the exogenous and idiographic methods, the window is not shown at all in the current developed version of GIB. Another improvement that was implemented for the dialog box was to list the methods by group (serial, idiographic and exogenous), hopefully leading the user to understand that the groups of methods differ principally. Additionally, in order to show the ideal distributional form for the serial types of classification methods, GIB was equipped with a ranked goodness, which applied both to serial and idiographic methods for class interval selection. The method having the highest ranking becomes the default option (see Figure 4.3). The ranking is based on an index called *goodness of variance fit* (GVF) whose values will lie between zero and one. Higher GVF values correspond with more accurate classified maps (see Rød, 2000; VI).

4.6 Towards transparent statistical mapping

Transparency might be included in statistical mapping if the resulting maps were attached with indications of their classification accuracy indicating “where” the maps hide significant differences or exaggerate insignificant ones. The techniques are outlined more profoundly in Rød (2000) and Rød (VI). Although these techniques presumably contribute to more accurate statistical maps, biases may still arise due to differences in the size and form of the administrative units used for statistical mapping. This is well known among cartographers from various empirical results (Dykes, 1994: 105) and is a motivation for using a regular grid zonal system (Bracken, 1994: 81). In *How the monosemic graphics go polysemic* (Rød, 2001a) this problem is treated in a theoretical manner by applying semiotic theory from a Saussurean tradition. Bertin, who without making any reference uses much of the same terminology as introduced by Saussure, elaborates the relations between visual marks, but overlook two other sign relationships: The relation between the visual marks and the geographical units wherein it is situated, and the relation between geographical units. By regarding the statistical geographical zones as Schlichtmann (1990: 265) does, as localized signs, the relationships between these make up an additional code in the cartographically transcribed meaning. While the visual variation between marks makes up the denotative semiotic, the visual variation between the geographical zones makes up a connotative semiotic (see Rød, 2001a, Figure 9). Since it is only the relationship of the former variation that is a priori defined[3], Bertin’s claim that graphics represent a monosemic sign system is consequently rejected regarding maps. A suggested practical solution might be to replace the more or less arbitrary administrative units with varying size and form with grid cells with equal size and form, in the same manner used in the census atlas of people in Britain (Clarke et al, 1980) and the population atlas from Sweden (Öberg and Springfeldt, 1991), and as also recently attempted in the mapping settlement patterns for Oslo (Rogstad, 2001: 27).

NOTES

[1] Checked after each hundredth iteration.

[2] Instead of ‘graph array’, ‘sorted diagram’ is used in earlier text describing this functionality.

[3] This is how Bertin defined a monosemic sign system: ‘A system is monosemic when the meaning of each sign is known **prior to** observation of the collection of signs’. By contrast ‘a system is polysemic when the meaning of the individual sign **follows** and is deduced from consideration of the collection of signs’ (Bertin, 1983, 2 – originally emphasis). See also Rød (2001a).

References:

Andrews J.H. 1996. What was a map? The lexicographers reply. *Cartographica*. 33 (4): 1-11.

Andrienko, G.L. and Andrienko, N.V. 1999. Interactive maps for visual data exploration. *International Journal of Geographical Information Science.* 13(4): 355-374.

Baudouin, A. 1987. Klasseinndeling i kartografi. *Kart og Plan.* 47(3) : 321-326.

Belyea, B. 1992. Images of power: Derrida/Foucault/Harley. *Cartographica*. 29 (2): 1-9.

Benveniste, E. 1971. The nature of the linguistic sign. In: *Problems in General Linguistics.* Benveniste, E. (Ed). Coral Gables, Florida: University of Miami Press. 43-49.

Bertin, J. 1967. *Sémiologie Graphique: Les Diagrammes, les Réseaux, les Cartes.* Paris: Mouton.

Bertin, J. 1977. *La Graphique et le Traitement Graphique de l'Information.* Paris: Flammarion.

Bertin, J. 1981. *Graphics and Graphic Information-Processing.* Translated by Berg, W.J. and Scott, P. Berlin: Walter de Gruyter.

Bertin, J. 1983. *Semiology of Graphics. Diagrams, Networks, Maps.* Translated by Berg, W.J. Madison, Wis.: University of Wisconsin Press.

Bertin, J. 1983b. A new look at cartography. In: *Graphic Communication and Design in Contemporary Cartography*. Taylor, D.R.F. (Ed). Chichester: Wiley. 69-86.

Blakemore, M. 1981. From way-finding to map-making: the spatial information fields of aboriginal peoples. *Progress in Human Geography.* 5 (1): 1-24.

Board, C. 1981. Cartographic communication. *Cartographica.* 18 (2): 42-78.

Bracken, I. 1994. Towards improved visualization of socio-economic data. In: *Visualization in Geographical Information Systems*. Hearnshaw, H.M. and Unwin, D.J. (Eds). Chichester: Wiley. 76-84.

Buttenfield, B.P. and Mackaness, W.A. 1991. Visualization. In: *Geographical Information System: Principles and Applications.* Maguire, D.J. Goodchild, M.F. Rhind, D. (Eds). Harlow: Longman. 427-443.

Chrisman, N.R. 1987. Design of geographical information systems based on social and cultural goals. *Photogrammetric Engineering and Remote Sensing*. 53 (10): 1367-1370.

Clarke, J.I., Dewdney, J.C., Evans, I.S., Rhind, D.W., Visvalingam, M., Coulter, J., Perry, B., Stanness, K., Denham, C. 1980. *People in Britain. A Census Atlas*. London: Her Majesty's stationery office.

Cromley, R.G. 1995. Classed versus unclassed choropleth maps: A question of how many classes. *Cartographica*. 32 (4): 15-27.

Curry, M.R. 1995. Geographical information systems and the inevitability of ethical inconsistency. In: *Ground Truth. The Social Implications of Geographical Information Systems*. Pickles, J. (Ed). New York: Guilford Press. 68-87.

Derrida, J. 1997. *Of Grammatology*. Translated by Gayatri Chakravorty Spivak. Corrected ed. Baltimore: Johns Hopkins University Press.

DiBiase, D. 1990. Visualization in the earth sciences. *Earth and Mineral Sciences*. 59 (2): 13-18.

Duncan, J. 2000. Text. In: *Dictionary of Human Geography*. Johnson, R.J., Gregory, D., Pratt, G. and Watts, M. (Eds). 4th ed. Oxford: Blackwell. 824-826.

Dykes, J. 1994. Area-value data: New visual emphases and representations. In: *Visualisation in Geographical Information Systems*. Hearnshaw, H.M. and Unwin, D.J. (Eds). Chichester: Wiley. 103-114.

Egbert, S.L. and Slocum, T.A. 1992. EXPLOREMAP: An exploration system for choropleth maps. *Annals of the Association of American Geographers*. 82 (2): 275-288.

Evans, I.S. 1977. The selection of class intervals. *The Institute of British geographers. Transactions. New Series*. 2: 98-124.

Forrest, D. 1993. Expert systems and cartographic design. *Cartographic Journal*. 30 (2): 143-148.

Godlewska, A. 1989. To surf or to swim. In: Commentary. Dahl, E.H. (Ed). *Cartographica*. 26 (3 & 4): 89-127.

Gregory, D. 1994. Discourse. In: *The Dictionary of Human Geography*. Johnston, R.J., Gregory, D. and Smith, D.M. (Eds). 3rd ed. 136-137.

Harley, J.B. 1989. Deconstructing the map. *Cartographica*. 26 (2): 1-20.

Harley, J.B. 1990. Cartography, ethics and social theory. *Cartographica*. 27 (2): 1-23.

Helbig, G. 1974. *Geschichte der Neueren Sprachwissenschaft*. Reinbek: Rowohlt.

Hirschheim, R., Klein, H.K. and Lyytinen, K. 1995. *Information Systems Development and Data Modelling. Conceptual and Philosophical Foundations*. Cambridge: Cambridge University Press.

Jenks, G.F. 1967. The data model concept in statistical mapping. *International Yearbook of Cartography*. 7: 186-190.

Jenks, G.F. Caspall, F C. 1971. Error on choroplethic maps: definition, measurement, reduction. *Annals of the Association of American Geographers*. 61 (2): 217-244.

Keates, J.S. 1996. *Understanding Maps*. 2nd ed. Harlow: Longman.

Kennedy, S. 1994. Unclassed choropleth maps revisited / Some guidelines for the construction of unclassed and classed choropleth maps. *Cartographica*. 31(1): 16-25.

Kuhn, T.S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kraak, M.J. and Ormeling F.J. 1996. *Cartography – Visualisation of Spatial Data*. Harlow: Longman.

Lang, L. 1995. The democratization of GIS: Bringing mapping to the masses. *GIS World* 8 (4): 62-67.

Langer, S.K. 1951. *Philosophy in a New Key*. New York: New American Library.

MacEachren, A.M. 1995. *How Maps Work. Representation, Visualization, and Design*. New York: Guilford Press.

MacEachren, A.M. and Ganter, J.H. 1990: A pattern identification approach to cartographic visualization. *Cartographica*. 27 (2): 64-81.

McGuinnes, C. 1994. Expert/novice use of visualization tools. In: *Visualization in Modern Cartography*. MacEachren, A.M. Taylor, D.R.F. (Eds). Oxford: Pergamon. 185-199.

McHaffie, P., Andrews, S.K., Dobson, M. and two anonymous employees of a federal mapping agency. 1990. Ethical problems in cartography. *Cartographic Perspectives* 7: 3-13.

McKenzie, D.F. 1986. *Bibliography and the Sociology of Texts*. London: British Library.

Miller, R.P. 1992. Beyond method, beyond ethics: Integrating social theory into GIS and GIS into social theory. GIS/LIS Annual Conference and Exposition, November 10-12, 1992, San Jose, California. 585-593.

Monmonier, M. 1991. Ethics and map design. Six strategies for confronting the traditional one-map solution. *Cartographic Perspectives* 10: 3-8.

Monmonier, M. 1991b. *How to Lie with Maps*. Chicago: University of Chicago Press.

Morris, C.W. 1938. *Foundations of the Unity of Science: Towards an International Encyclopaedia of Unified Science, Volumes 1 and 2*. Chicago: University of Chicago Press.

Muehrcke, P.C. 1978. *Map Use: Reading, Analysis, and Interpretation.* Madison: JP Publications.

Nöth, W. 1990. *Handbook of Semiotics.* Bloomington: Indiana University Press.

Müller, J.C. and Wang, Z. 1990. A knowledge based system for cartographic symbol design. *Cartographic Journal.* 27 (1): 24-30.

Ormeling, F.J. 1983. *Minority names on maps: the rendering of minority toponyms on topographic maps of Western Europe.* Ph.D. thesis. Utrecht: Series Utrechtse Geografische Studies no. 30.

Ormeling, F.J. 1992. Brian Harley’s influence on modern cartography. *Cartography*. 29(2): 62-66.

Papp-Váry, A. 1989. The science of cartography, In: *Cartography Past, Present and Future.* Rhind, D.W. and Taylor, D.R.F. (Eds). London, New York: Published on behalf of the ICA by Elsevier Applied Science. 103-109.

Robinson, A.H. and Petchenik, B.B. 1976. *The Nature of Maps - Essays Toward Understanding Maps and Mapping.* Chicago: University of Chicago Press.

Robinson, A.H., Sale, R.D., Morrison J.L., Muehrcke P.C. 1984. *Elements of Cartography*. New York: Wiley. 5th edition.

Robinson, A.H., Morrison J.L., Muehrcke P.C., Kimerling A.J., Guptill S.C. 1995. *Elements of Cartography.* New York: Wiley. 6th edition.

Rogstad, L. 2001. Rutenettbasert statistikk, belyst ved bosettingsmønsteret i Oslo. *Geodata Magasinet*. Vol. 13. February.

Rundstrom, R.A. 1993. The role of ethics, mapping, and the meaning of place in relations between Indians and Whites in the United States. *Cartographica.* 30(1): 21-28.

Russell, B. 1993. *Our Knowledge of the External World: As a Field for Scientific Method in Philosophy.* London: Routledge. Originally published: Open Court Pub. Co., 1914.

Rød, J.K. 1998. Le troisième choix ... *Comité Français de Cartographie.* 156: 98-102.

Rød, J.K. 1999. Vurdering av funksjonalitet ved fremstilling av koropletkart i Excel. *Kart og Plan*. 59 (1): 36-42.

Rød, J.K. 2000. De nauwkeurigheid van geclassificeerde kaarten. *Kartografisch Tijdscrift* 16 (4): 47-51.

Rød, J.K., Ormeling, F.J. and Van Elzakker, C.P.J.M. 2001. An agenda for democratising cartographic visualisation. *Norwegian Journal of Geography*. 55 (1): 38-41.

Rød, J.K. 2001a. How the monosemic graphics go polysemic. *Cartographic Perspectives.* <in press>.

Rød, J.K. 2001b. Cartography, the real, and the good. In: *Contemporary Debates in the Discipline of Geography. Space and Place. Landscape and Environment*. Lægran, A.S. and Myklebust, F. (Eds). Report from a Research Course in Geography, Bårdshaug, September 2000. Papers from the department of Geography. New Series A. 16: 157-170.

Rød, J.K. VI. The selection of class intervals revisited. Submitted to *Transactions of the Institute of British Geographers*.

Sack, R.D. 1999. A sketch of a geographic theory of morality. *Annals of the Association of American Geographers.* 89 (1): 26-44.

Saussure, F. 1974. *Course in General Linguistics*. Rev. ed. London: Owen.

Schlichtmann, H. 1990. Plan information and its retrieval in map interpretation: the new view from semiotics. In: *Cognitive and linguistic aspects of geographic space.* Mark, D.M. Frank, A.U. (Eds). Dordrecht: Kluwer Academic Publishers. 263-284.

Smith, N. 1992. History and philosophy of geography: real wars, theory wars. *Progress in Human Geography.* 16 (2): 257-271.

Someren, M.W. van, Barnard, Y.F. and Sandberg, J.A.C. 1994. *The Think Aloud Method. A Practical Guide to Modelling Cognitive Processes.* London: Academic Press.

Stevens, S.S. 1946 On the theory of scales of measurement. *Science*. 103 (2684): 677-680.

Sølvberg, A. and Kung, D.C. 1993. *Information Systems Engineering. An Introduction.* Berlin: Springer.

Tobler, W.R. 1973. Choropleth maps without class intervals? *Geographical Analysis* 5 (3): 262-264.

Wang, Z. and Ormeling, F.J. 1996. The representation of quantitative and ordinal information. *Cartographic Journal.* 33 (2): 87-91.

Weibel, R. and Buttenfield, B.P. 1992. Improvement of GIS graphics for analysis and decision-making. *International Journal of Geographical Information System.* 6 (3): 223-245.

Wood, D. 1992. *The Power of Maps.* New York: Guilford Press.

Öberg, S. and Springfeldt, P. 1991. *National Atlas of Sweden: The Population.* Stockholm: SNA

Part II

Articles

Article I

Rød, J.K., Ormeling, F.J. and Van Elzakker, C.P.J.M. 2001. An agenda for democratising cartographic visualisation. *Norwegian Journal of Geography*. 55 (1): 38-41.

Norsk geogr. Tidsskr. Vol. 55, 38–41. Oslo. ISSN 0029-1951

An agenda for democratising cartographic visualisation

JAN KETIL RØD, FERJAN ORMELING & CORNÉ VAN ELZAKKER

Rød, J. K., Ormeling, F. & Van Elzakker, C. P. J. M. 2001. An agenda for democratising cartographic visualisation. *Norsk Geografisk Tidsskrift–Norwegian Journal of Geography* Vol. 55, 38–41. Oslo. ISSN 0029–1951.

Cartographic visualisation tools aid exploration, but they are designed for, and used exclusively by, experts. A democratised visualisation tool will include second-generation users, and these non-specialists might also want to use the available computer technology to visualise their geographical data. In this paper, we argue that democratised GIS should have a functionality similar to visualisation tools and we forward our opinion on how these can be developed in order to do so. Our emphasis is on interactivity regarding representation methods and on elaborating principles for implementing map type selection in interfaces for democratised GISs.

Keywords: *exploratory cartography, interactive mapping systems, statistical mapping*

Jan Ketil Rød, Department of Geography, NTNU (Norwegian University of Technology and Science), N-7491 Trondheim, Norway. E-mail: jan.ketil.rod@svt.ntnu.no; Ferjan J. Ormeling, Department of Cartography, Utrecht University, P.O. Box 80115, 3508 TC Utrecht, The Netherlands. E-mail: f.ormeling@geog.uu.nl; Corné P. J. M. Van Elzakker, Division of Geoinformatics, Cartography and Visualization, International Institute for Aerospace Survey and Earth Sciences (ITC), P.O. Box 6, 7500 AA Enschede, The Netherlands, E-mail: elzakker@itc.nl

Introduction

For several years now, users of standard software for spreadsheets have been able to present their geographical tabular data as maps. MS Excel, for instance, has embedded a simplified version of MapInfo, an event commented on in *GIS World* by Lang (1995) in an article on the democratisation of GIS. Democratisation is a concept used for describing a situation where disciplines once highly specialised become available for nearly everyone; since 1995 this has been applied to cartography (Dorling & Fairbairn 1997, Morrison 1997). Below, we denote standard software that has embedded mapping facilities as democratised cartographic software or democratised GIS. As spreadsheet packages are more or less standard equipment for offices, according to Lang (1995, 62), anyone among some 50 million users is able to create their own choropleth maps, dot density maps, proportional symbol maps, chorochromatic maps, bar chart maps, or pie chart maps. These types of maps might all be termed statistical maps, aiming to portray the geographical distribution of one or several variables. Essential for the manner in which a geographical distribution is portrayed, however, is the type of representation method selected and the design principles applied. Perhaps the democratisation of GIS will lead to an increased understanding of the manner in which cartographic representation and design options influence the portrayal of geographical distributions. The attitudes of professional cartographers seem contrary to this trend, so far.

Expert systems and experienced learning

A view asserted by several is that if everyone becomes a mapmaker this will result in unsound maps (Müller & Wang 1990, 24, Weibel & Buttenfield 1992, 223, Forrest 1993, 144). Consequently, considerable concern has been directed towards developing expert systems as part of mapping software in order to guarantee the production of proper maps. If the users themselves do not have the necessary cartographic knowledge, it was argued, this knowledge ought to be available within the GIS or cartographic packages. Many of the approaches to expert systems in cartography were, and still are, restricted to the embedding of cartographic knowledge into smaller, clearly bound domains such as map design (McMaster & Buttenfield 1997, 207). Forrest (1999), however, has presented a holistic approach to a cartographic expert system. Previously, such holistic approaches were rare, probably because cartography as a discipline was considered too large to fit a single knowledge-based system and therefore difficult to lay down in a set of rules (Kraak & Ormeling 1996, 210). In this paper, we take an experienced learning perspective as our approach. The notions *expert* and *experience* both imply dealing with knowledge. Approaches to expert systems tend to concentrate on the necessity of embedding expert knowledge, while an experienced learning perspective concentrates on the possibilities for anybody to gain new knowledge. We emphasise the latter and we want to set an agenda for enhancing the ordinary user's possibilities of using maps or map displays as tools to explore geographic data. In this respect, an understanding of the influence that representation methods (the selected types of maps) and design principles have on the portrayed distribution is essential. If the users themselves do not have the necessary cartographic knowledge, we argue, this knowledge might be obtained provided there are resources and appropriate program responses available for the users.

Cartographic communication versus cartographic visualisation

We adopt a key distinction between maps fostering *private visual thinking* directed towards revealing unknowns (exploration) and maps facilitating *public visual communication* directed towards presenting knowns (presentation) (DiBiase 1990, 14). As the term *visualisation* in the title indicates, our main emphasis is on the exploratory use of maps. Usually, visualisation is denoted with the prefix *scientific*, defined as 'the use of sophisticated computer technology to create visual displays, the goal of which is to facilitate thinking and problem solving' (Kraak & Ormeling 1996, 198). Therefore, advancements in cartographic visualisation tend to direct themselves towards highly specialised user groups: 'we assume a person doing visualization process on a computer is a professional, often with an academic education. He or she is working with dedicated software on some specialized problem on which he or she is an expert' (Lindholm & Sarjakoski 1994, 173). Less attention has been paid towards what we denote as democratised cartographic visualisation. As GISs and cartographic packages become increasingly generally available:

> no longer does the map user depend upon what the cartographer decides to put on a map. Today, the user is a cartographer. This represents a "democratisation" of cartography in which all individuals are potentially empowered with the available electronic tools to think geographically and to make visualisations of their thinking. (Morrison 1997, 17)

Cartography is perhaps democratised in the meaning that everybody can become a mapmaker, but we claim that cartographic visualisation is not democratised at all, since these software packages inadequately facilitate geographic thinking and the visualisation of these thoughts. We believe, like Young, that a 'failure to make and use maps' and map displays 'in a meaningful way contributes to a lack of geographic awareness' (Young 1994, 10). In order for democratised cartographic software packages to function as tools to reveal unknown geographies, the degree of human–map interaction is recognised as an essential factor. 'For cartographic visualization tools to succeed interaction is paramount; the system should permit the user to do a wide variety of things to the data' (MacEachren & Ganter 1990, 74).

There are several examples of interactive cartographic systems. Exploremap (Egbert & Slocum 1992) enables the user to explore univariate attribute data associated with choropleth maps. In the impressive interactive cartographic program Descartes (Andrienko & Andrienko 1998), not only choropleth maps but also maps portraying nominal data, proportional symbol maps, and multivariate maps can be used to explore the data. The users of Descartes can easily study patterns of spatial relationships by manipulating the design principles. As a rule in the Descartes program multiple presentations of the same data are complementary, since each presentation method suits certain types of analysis tasks. It is important, however, not only to relate this to *presentation* method, i.e. to design principles, but also to *representation* method, i.e. map types. While several articles direct themselves towards issues of design principles (Forrest 1993, McMaster & Buttenfield 1997), less attention has been directed to issues of representation methods. Its importance is addressed by Keates, who stated that 'misjudgements in the choice of representation methods require far more attention in relation to statistical maps than they normally receive' (Keates 1996, 306). We would like the users of democratised GISs to understand the implications various representation methods have on the portrayal of geographical distributions. The users might very well not be aware of all cartographic methods of representation available for geographical data exploration. Democratised cartographic visualisation tools should therefore focus on the way different representation methods might suit various exploratory tasks.

What do you want to see?

The preparatory stage for making maps in Excel is completely automated. The mapping situation can be illustrated with the through-going grey, thick arrow in Fig. 1. From the spreadsheet, the users select the data holding the units and variable(s) they want to see mapped. Automatically, a choropleth map shows the variable classified according to the equal interval principle with five classes of graded shades of grey. This releases the users from the work of selecting representation method and map design. Ideally, the users would therefore be able to devote more time and attention to data analysis. However, since each representation method suits certain types of analysis tasks better than others, this might easily lead to an unsuitable image. The map control in Excel is flexible in terms of allowing the users to reselect the variable to be mapped, reselect representation method, and reselect design principles (indicated with the returning arrows in Fig. 1). However, the mapping situation could be improved, first by offering a default representation method based on some argument. Second, the users should be able to reflect on which of the representation methods offered suits a particular task.

The latter could be realised by implementing a truly interactive mapping system. 'Truly interactive mapping systems should present the user with a number of choices, options and decisions including (1) on-screen choices, (2) resources to make these choices and (3) appropriate program responses' (Asche & Herrmann 1994, 220). Just as users of the statistical coach in SPSS (Statistical Package for the Social Sciences) are asked: *What do you want to do?* – users of an interactive mapping tool can be asked: *What do you want to see?* The users may, for instance, want to look for differences between areas, a general pattern or dominance of one variable related to another. Several representation options should then be available, but as different types of maps show different types of information, explanations on what the users might expect to see by selecting a specified map will be useful. In Fig. 2, the main types of maps that are, should, or should not be available in the mapping module of a software package such as MS Excel are presented. Note that we distinguish between the dot density map, as available in the Excel module, and the dot map we would prefer. A dot map is a map using 'a series of uniform point symbols to represent a quantity of data by the repetition of a point symbol. Each point symbol is equated to so much of the distribution being mapped, and its placement makes an attempt to show the location of the distribution' (Robinson et al. 1984, 300). The dot density map on the other hand, constitutes another map type since it does not show the locational character of a distribution, nor is this, according to manuals, the intention. Dots are located randomly in dot density maps and do not provide useful information regarding concentration or dispersion or other pattern characteristics. Consequently, we do not advocate the use of dot density maps.

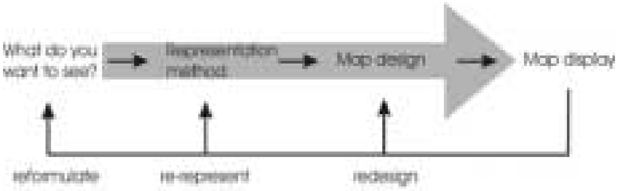


Fig. 1. The visualisation process

Neither do we advocate the use of pie chart and bar chart maps for bi- or multivariate mapping. In order to see dominance, we find it better to map two variables, e.g., by using 'an ellipse as a map symbol in which the horizontal and vertical axes represent different (but presumably related) variables' (MacEachren 1995, 90).

Regarding default options, we argue that, if offered, default options of representation methods should be based on some argument. This argument is related first of all to the nature of the geographical data portrayed. It should not only be known *where* the data are (location), but also *what* the data are (characteristics). In this latter context, two complementary instruments could support these default options: the characteristics of the data and base map dimensionality.

The characteristics of data might be classified as shown in Fig. 3. The data may be either qualitative or quantitative, and, if quantitative, they may be either numerical or not, and if numerical, they may be either absolute or relative. The characteristics of the data determine the adequate cartographic representation that a specific variable may take. Having this information structured as metadata[1] would therefore be profitable in order to advise map type selection. Similar approaches are present in closed systems such as the Atlas Map (Forrest 1995), Descartes (Andrienko & Andrienko 1998), and Statistisk Sett[2] packages. These are closed systems since both maps and attribute data constitute an embedded part of the program and are never opened or imported. Consequently, it is easy to attach the variables with metadata in order to apply constraining principles. For instance, the beta version of the Statistisk Sett package prevented users making choropleth maps from absolute ratio values. Instead, a message informed them that choropleth mapping should only be applied for relative values. In the commercially available version, the users do not have the options of making choropleth maps from absolute values. The variables have either relative or absolute ratio values and their representation methods are choropleth maps or proportional point symbol maps respectively.

Recent developments are also approaching an embedding of constraining principles based on the characteristics of the data for open systems. When a new variable is to be defined in SPSS, the level of measurement is among the parameters describing the variable. The users might accept, in ignorance or otherwise, the default option *scale* (numeric data on an interval or ratio scale), or select one of the two other options *ordinal* or *nominal*. When importing SPSS-format data files created in earlier versions of SPSS products, rules are applied in order to label the variable with information on measurement level. Although this leads to errors,[3] it is a genuine start to solving a complex problem.

Both in MapInfo and in the mapping module of Excel, only six representation methods are available. This is mainly because, in these packages, only one approach to the representation of statistical data is implemented. There are five approaches that cartographic software geared towards the representation of statistical data might take:

1. An individual-level where data relate to an individual entity in a population (point representation).
2. An assumption that the phenomena of interest flow *along* lines or *across* lines (linear representation).
3. An area aggregation – administrative areas or grid squares – of individual-level data (area representation).
4. An assumption that the phenomena of interest are essentially continuous over space, with attempts to reconstruct this spatial continuity (surface representation).
5. An assumption that the phenomena of interest are essentially continuous over time, with attempts to reconstruct this temporal continuity (temporal representation).

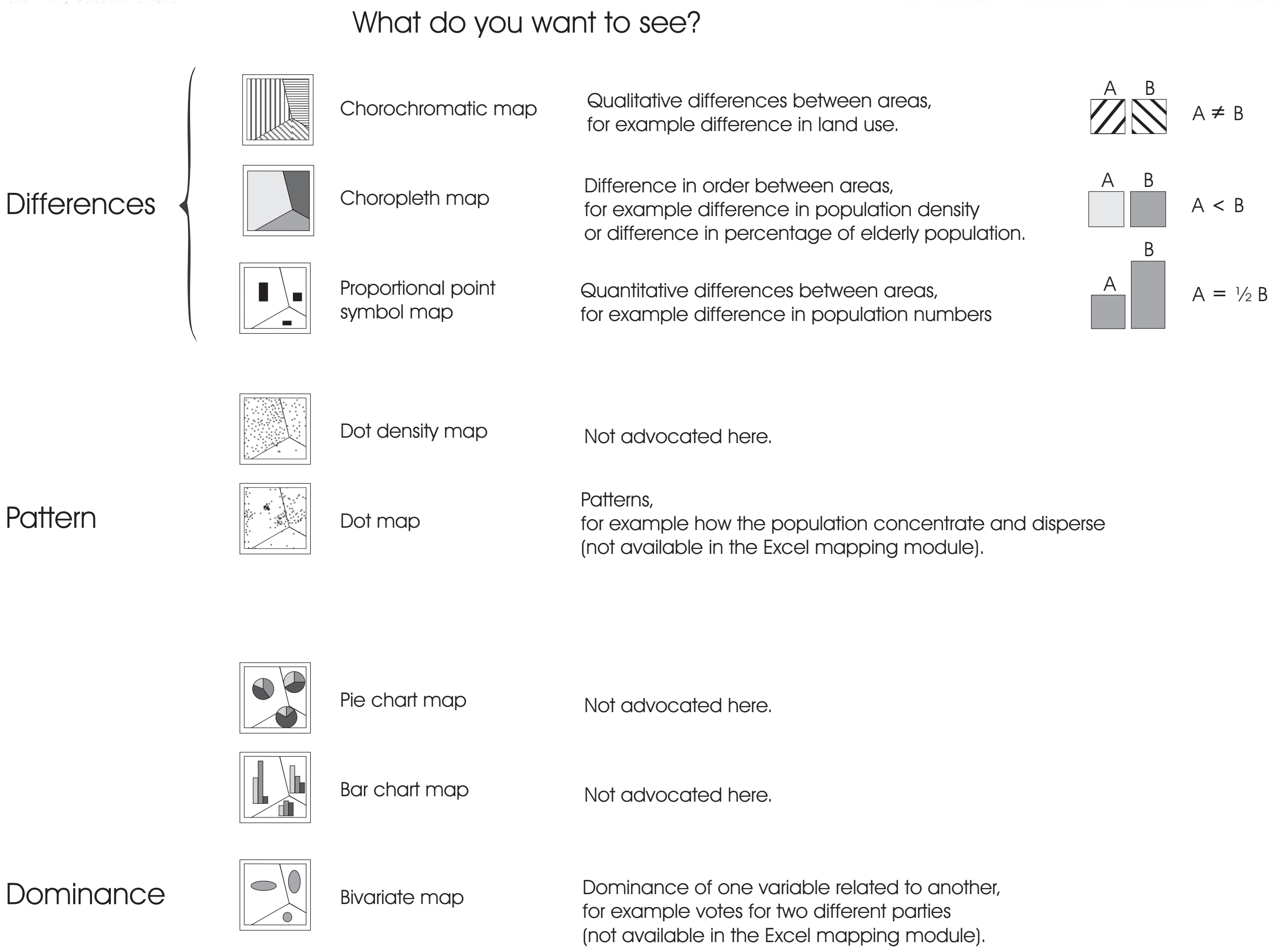


Fig. 2. Elements for an interactive mapping system.

While a GIS such as ArcView with the extensions Spatial Analyst, 3D Analyst and Tracking Analyst covers all of these approaches, the current approach for democratised cartographic software is the area aggregation approach only.

In Table 1, point, linear and area aggregation approaches are combined, resulting in 12 different univariate map types. Nine are already available in ArcView, but only six in the Excel mapping module. We argue that a program for democratised cartographic visualisation should try to combine several of the above approaches because the representation methods they provide are also very important in an exploratory environment. By using the characteristics of the data and the representational character (e.g. point) as a combined constraining factor, Table 1 might be a point of departure for implementing a support for the selection of statistical map types. For example, if the variable to be mapped holds nominal data related to areas, only the chorochromatic map option will be available.

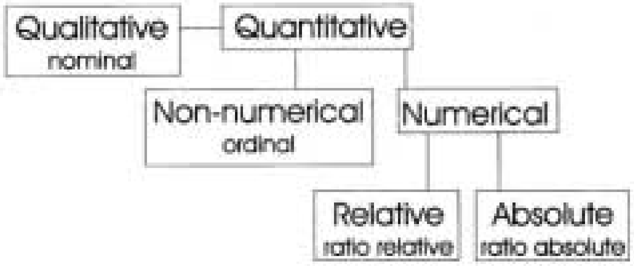


Fig. 3. Characteristics of the data.

Concluding remarks

Cartographic visualisation tools exist mainly in the domain denoted *scientific*, but hardly, if at all, in a domain we might denote *public*. An extension, however, is likely to take place.

> In common with first-generation users of other computer technologies, many of the new visualization tools are designed for, and used exclusively by, experts. Although their application might be currently restricted to an expert group, second generation users are likely to include a wider community who may not be as adept at spatial reasoning as geographers, environment scientists, surveyors and cartographers. Increasingly the new tools will be used in education and to communicate with, and persuade, colleagues, policy-makers, planners and the public. (McGuinness 1994, 197)

Instead of viewing the extension of the group of mapmakers as menacing proper mapmaking, we believe in another possible outcome of the democratisation of GIS. If democratised GISs function as visualisation tools, this might lead to an increased general understanding of how maps and map displays function both as tools for exploring geographic phenomena and as devices for communicating knowledge about these phenomena.

The first point relates to how new conceptions about our environment might evolve. Regardless of whether the user doing spatial reasoning is an expert or novice, interactivity is paramount for the discovery of geographical patterns. Adding interactivity rather than reducing this by stripping choice

Table 1. Constraining metadata information for map type selection.

Characteristics of the data	Base map dimensionality	Corresponding map type
Nominal	Point	Dot map or point symbol map for nominal data
	Line	Line symbol map for nominal data
	Polygon	Chorochromatic map
Ratio–relative or ordinal	Point	Ordered point symbol map
	Line	Ordered line symbol map
	Polygon	Choropleth map
Ratio–absolute	Point	Proportional point symbol map
	Line	Flow map (proportional line symbol map)
	Polygon	Cartograms (proportional area symbol map)
Ratio–absolute	Point	Dot map
	Line	Flow dot map (dots located at lines)
	Polygon	Dot density map

options and flexibility will attain user-friendliness. We believe, like Raper (1991, 111), that 'gaining experience with the alternative options is an excellent way to improve a user's end-to-end understanding of the components of spatial data processing'.

The second point relates to a situation where maps and mapping become more general among users expressing conceptions about our environment. People are generally less sceptical towards maps than to written or oral forms of expressions (Monmonier 1995, 1). However, as these modes of expression become available for everyone, their manipulating power is likely to decline because of increasing suspiciousness. If the democratisation of GIS 'demystifies the making of maps' (Dorling & Fairbairn 1997, 128), the general view of the map might turn to an expressed perspective about concepts of the world. Map expressions will then meet their critical voices in the same way as assertions in textual or oral forms.

Manuscript submitted 5 March 2000; accepted 31 October 2000.

Notes

[1] Metadata are additional information prepared by the data producer and consist of a set of structured digital information that describes the data. 'A set of metadata allows the data user to judge the information's fitness for a given application' (Robinson et al. 1995, 197).

[2] *Statistisk Sett* is cartographic software produced by the Norwegian Mapping Authority, the Norwegian Statistical Bureau and the ESRI vendors in Norway: Geodata.

[3] For instance, when a variable holding sex is coded 1 for 'male' and 2 for 'female', this nominal variable is defined as being ordinal.

References

Andrienko, G. L. & Andrienko, N. V. 1998. Interactive maps for visual data exploration. The ICA commission on visualization meeting (Warsaw, Poland, May 21st–24th 1998), http://borneo.gmd.de/and/icavis/, July 18, 2000.

Asche, H. & Herrmann, C. M. 1994. Designing interactive maps for planning and education. MacEachren, A. M. & Taylor, D. R. F. (eds.) *Visualization in Modern Cartography*, 215–242. Elsevier, Oxford.

DiBiase, D. 1990. Visualization in the earth sciences. *Earth and Mineral Sciences* 59 (2), 13–18.

Dorling, D. & Fairbairn, D. 1997. *Mapping: Ways of Representing the World.* Longman, Harlow.

Egbert, S. L. & Slocum, T. A. 1992. Exploremap: an exploration system for choropleth maps. *Annals of the Association of American Geographers* 82 (2), 275–288.

Forrest, D. 1993. Expert systems and cartographic design. *Cartographic Journal* 30 (2), 143–148.

Forrest, D. 1995. Don't break the rules on helping non-cartographers to design maps: an application for cartographic expert systems. *17th International Cartographic Conference, Barcelona, September 3rd–9th 1995*, 1, 570–579.

Forrest, D. 1999. Developing rules for map design: a functional specification for a cartographic design expert system. *Cartographica* 36:3, 31–52.

Keates, J. S. 1996. *Understanding Maps.* 2nd ed. Longman, Harlow.

Kraak, M. J. & Ormeling, F. J. 1996. *Cartography: Visualization of Spatial Data.* Longman, Harlow.

Lang, L. 1995. The democratization of GIS: bringing mapping to the masses. *GIS World* 8 (4), 62–67.

Lindholm, M. & Sarjakoski, T. 1994. Designing a visualization user interface. MacEachren, A. M. & Taylor, D. R. F. (eds.), *Visualization in Modern Cartography,* 167–184. Elsevier, Oxford.

MacEachren, A. M. 1995. *How Maps Work: Representation, Visualization and Design.* Guilford Press, New York.

MacEachren, A. M. & Ganter, J. H. 1990. A pattern identification approach to cartographic visualization. *Cartographica* 27 (2), 64–81.

McGuinness, C. 1994. Expert/novice use of visualization tools. MacEachren, A. M. & Taylor, D. R. F. (eds.) *Visualization in Modern Cartography,* 185–199. Elsevier, Oxford.

McMaster, R. & Buttenfield, B. 1997. Formalising cartographic knowledge. Craglia, M. & Conclelis, H. (eds.) *Geographic Information Research: Bridging the Atlantic,* 205–223. Taylor & Francis, London.

Monmonier, M. 1995. *Drawing the Line: Tales of Maps and Cartocontroversy.* Henry Holt, New York.

Morrison, J. L. 1997. Topographic mapping in the twenty-first century. Rhind, D. (ed.) *Framework for the World,* 14–27. GeoInformation International, Cambridge.

Müller, J. C. & Wang, Z. 1990. A knowledge based system for cartographic symbol design. *Cartographic Journal* 27 (1), 24–30.

Raper, J. F. 1991. User interfaces. Masser, I. & Blakemore, M. (eds.) *Handling Geographical Information: Methodology and Potential Applications,* 102–114. Longman, Oxford.

Robinson, A. H., Sale, R. D., Morrison, J. L. & Muehrcke, P. C. 1984. *Elements of Cartography.* 5th ed. Wiley, New York.

Robinson, A. H., Morrison, J. L., Sale, R. D., Muehrcke, P. C., Kimerling, J. & Guptill, S. C. 1995. *Elements of Cartography.* 6th ed. Wiley, New York.

Weibel, R. & Buttenfield, B. P. 1992. Improvement of GIS graphics for analysis and decision-making. *International Journal of Geographical Information Systems* 6 (3), 223–245.

Young, J. E. 1994. Reexamining the role of maps in geographic education: images, analysis, and evaluation. *Cartographic Perspectives* 17, 10–20.

Article II

Rød, J.K. 1999. Vurdering av funksjonalitet ved fremstilling av koropletkart i Excel. *Kart og Plan*. 59 (1): 36-42.

Assessing the functionality of choropleth mapping using Excel.

Critical comments

Vitenskapelig bedømt artikkel.

Vurdering av funksjonalitet ved fremstilling av koropletkart i Excel

Jan Ketil Rød

Mottatt 02.11.1998. Godkjent 24.02.1999

Jan Ketil Rød: Assessing the Functionality of Choropleth Mapping using Excel.

KART OG PLAN, Vol. 59, pp. 36–42, P.O.B. 5034, N-1432 Ås, ISSN 0047-3278

For several years users of standard software for spreadsheets have been able to visualise their tabular data as maps due to a new tool: Statistical mapping. One such standard software that has imbedded mapping capabilities is Excel. An evaluation of Excel's functionality for statistical mapping is carried out. By functionality we mean the software's abilities to respond to the users intention. The evaluation deals with choropleth mapping. This mapping process is decomposed into three decisions: the choice of number of classes, the choice of method for class breaks determination, and choice of symbolisation. How the Excel mapping environment performs each of these steps is investigated in comparison with cartographic theory. Finally, some reflections on the influence the democratisation of mapping might have on map design is put forward.

Key words: Choropleth mapping. Map design. Visualisation. Democratisation.

Jan Ketil Rød, Research Fellow, Department of Geography, Norwegian University of Science and Technology, N-7491 Trondheim, Norway. E-mail: jan.ketil.rod@svt.ntnu.no

Innledning

Utgangspunktet for denne artikkelen er den allmenne tilgjengelighet til statistiske data samt muligheten for hvermann å fremstille disse data i kart. Det siste er muliggjort ved at alminnelig programvare for regneark og statistikk har fått en kartmodul inkludert. Brukere av for eksempel programvaren Excel kan nå selv lage koropletkart, prikkekart, størrelsesproporsjonale kart, kvalitative skravurkart, stolpe- eller kakediagramkart. Disse karttypene kommer alle inn under betegnelsen statistiske kart, hvis formål er å fremstille den *geografiske* fordeling av et tallmateriale – som for eksempel salgstall, demografiske data eller annet. Siden Excel og liknende programvare mer eller mindre er standard kontorutrustning, representerer dette en helt ny situasjon: hvem som helst blant omlag 50 millioner brukere av programvare for regneark, kan nå egenhendig produsere sine statistiske kart. Begivenheten ble kommentert i GIS world med artikkelen: *The Democratization of GIS – Bringing Mapping to the Masses* (Lang, 1995). Denne demokratiseringen, eller allminneliggjøringen av kartfremstilling, kan medføre produksjon av dårlige kart. Et annet mulig utfall er at kartografisk kunnskap om god kartdesign blir allminneliggjort. I følge Monmonier (1995:1) er kart mindre utsatt for kritisk vurdering enn hva tilfellet er for skriftlige former for kommunikasjon. At alle kan lage kart medfører kanskje en demystifisering av kartfremstilling (Dorling og Fairbain 1997:165) slik at synet på kartet som virkelighetsavbilding erstattes med synet på kartet som et argument, tilsvarende en påstand i tekstuell eller muntlig form. Da vil også antallet kritiske røster mot dårlig kartdesign øke.

Denne artikkelen begrenser seg til å vurdere funksjonaliteten til ett alminnelig EDB-verktøy hvor statistikk kan fremstilles i kart: Programvaren Excel. Med funksjonalitet menes her brukers mulighet til å gjøre valg og hvordan programmet responderer på disse. Kan brukers intensjoner utføres og får bruker noen form for tilbakemelding om konsekvenser av de valg hun gjør? Artikkelen vil i hovedsak avgrense seg til fremstilling av koropletkart der vurderinger om produksjon av tidsserie koropletkart er utelatt. Betraktninger om hvordan de administrative områdenes varierende størrelse påvirker koropletkartets visuelle bilde, er heller ikke inkludert i denne undersøkelsen.

En evaluering av koropletkart fremstilling i Excel

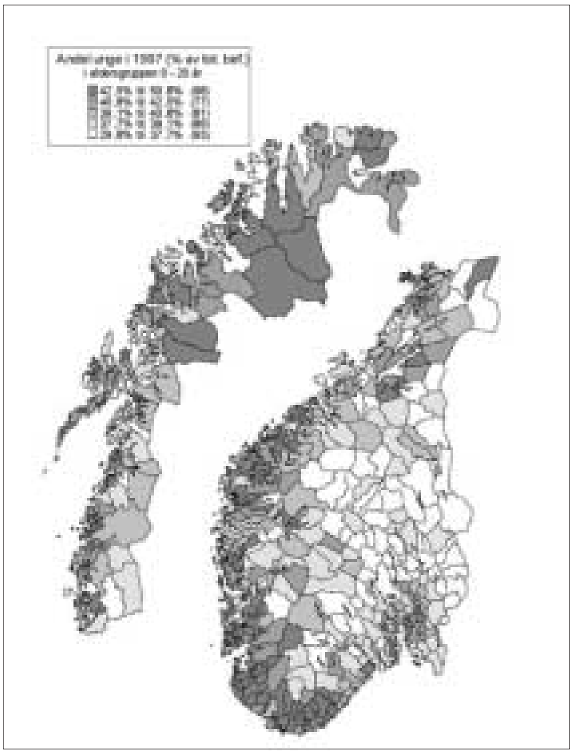
Når funksjonaliteten til koropletkart fremstilling i Excel skal evalueres, er det hensiktsmessig å dekomponere denne kartfremstillingsprosessen. Dette kan gjøres ved å atskille de tre avgjørelser som koropletkart produksjon i hovedsak omfatter (Baudouin 1987):

1. Valg av antall klasser
2. Valg av metode for å bestemme klassegrenser
3. Valg av grafisk fremstilling

På hvilket grunnlag og under hvilke betingelser tas disse tre avgjørelsene for en bruker av Excel programvare og hvordan blir resultatet? Spørsmålet skal i første omgang vurderes i forhold til Excels prototypekart, det standard kartet som dannes ved oppstart av kartmodulen, et kart som er dannet utelukkende på grunnlag av valg tatt av systemet. Deretter skal spørsmålet vurderes i henhold til brukers fleksibilitet med hensyn til å omgjøre disse tre beslutninger. Norsk Excel versjon 97 er benyttet i undersøkelsen.

Vurdering av standard kart design i Excel og brukers fleksibilitet

Figur 1 viser et eksempel på Excels standard kart (koropletkart med kvantil inndeling) av variabelen 'Andel unge (0 til 20 år) i 1997 i prosent av total befolkning' for norske kommuner[1] For å lage dette kartet må først variabelen og kolonnen med kommunenummer merkes i regnearket. Ved å klikke på ikonet som forestiller en globus, genereres deretter et kart. Dette gir som standard fremstilling et koropletkart med fem klasser. Klassegrensene er bestemt ved at klassene får tilnærmet like mange elementer i hvert område av verdier. Denne klassifikasjonsteknikk kalles vanligvis *percentil* eller *kvantil*[2] inndeling. Symboliseringen er en gradert gråtoneskala fra hvitt til mørk grå. Antallet observasjoner i hver klasse vises i parenteser i tegnforklaringen.



Figur 1: Excels prototypekart. Benyttet klasseinndeling er kvantil inndeling. Inndelingen forsterker en geografisk differensiering mellom norske kommuner.

Kartografisk praksis og forskning har resultert i en mer eller mindre konsensus om 'regler' for god kart design. Er Excels standard kart dannet i henhold til disse 'kartografiske regler'?

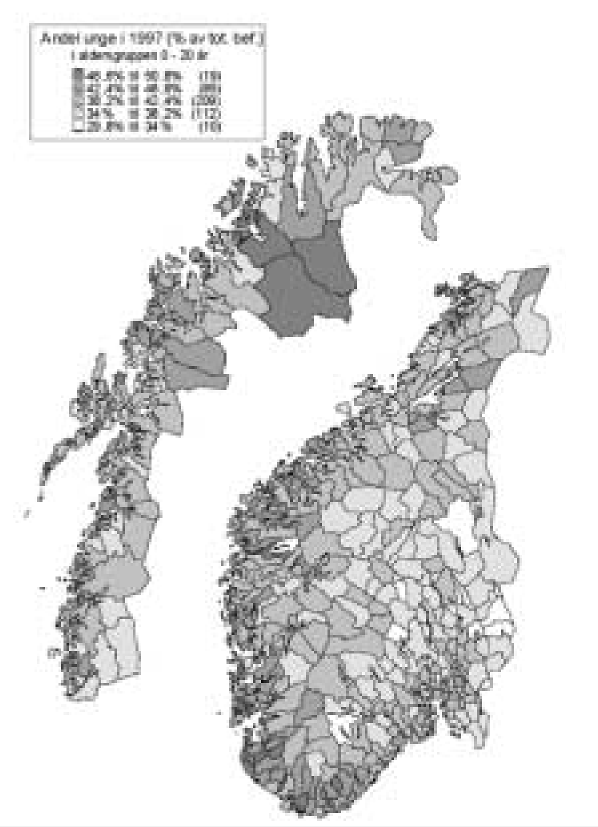
Valg av antall klasser

Siden vi kun er i stand til å skille mellom et begrenset antall gråtoner, begrunnes det ofte ut fra persepsjonshensyn at antall klasser bør begrenses. Som et kompromiss mellom ulike resultater og teorier er fem klasser vurdert som passende for en gråtoneskala mens antallet klasser kan økes noe om det benyttes en grade-

---

[1] Variablene som kartene i denne artikkelen fremstiller er avledet fra demografiske variabler i kommunedatabasen som leveres av Norsk Samfunnsvitenskapelig Datatjenestes (NSD). Koordinater for kommunenes grenser er også skaffet fra NSD. Basiskartene er generert i MapInfo, programmet som har levert kartmodulen til Excel, og deretter tilrettelagt for Excel.

[2] Kvantiler (quantile) er en mer generell term for percentiler. I engelsk språkbruk benyttes ofte percentiler for å unngå sammenblanding av 'quartiles' (kvartiler) og 'quintiles' (kvintiler).



Figur 2: Koropletkart med like intervall inndeling. Den regionale differensiering som var synlig i figur 1 er her sterkt dempet – norske kommuner ser like ut.

ring fra lyst til mørkt av en varm, kraftig farge (Mersey 1990). Som standard i Excel settes antall klasser til fem og benyttet symbolisering er en gradert serie i gråtoner. Moderprogrammet MapInfo benytter som standard syv klasser og en gradert symbolisering fra lys til mørk rødfarge[3]. Evans (1977) argumenterer at valg av antall klasser øker kartets kompleksitet og derfor bør valg av antall klasser også relateres til om brukerne er vant med å 'lese' grafisk informasjon eller ikke.

> Within the range four to ten classes, a decision should be influenced by the intended audience, the technical means available, and the spatial pattern of the distribution. A simple clear-cut map with four or five classes may be better for an unsophisticated audience, inexperienced at reading graphics. Trained eyes may appreciate the extra information which seven or eight classes portray (Evans 1977:100).

Brukere av MapInfo programvare vil nok være mer sofistikerte enn alminnelige brukere av Excel programvare som sannsynligvis vil komme i kategorien uerfarne med hensyn til kartografisk erfaring. Det er derfor fornuftig at standard antall klasser er færre i Excel enn i moderprogrammet MapInfo.

Med hensyn til valg av antall klasser foreslo Tobler (1973) en metode for klasseløse koropletkart. Det utløste den gang en debatt som i nyere tid er revitalisert av Kennedy (1994) og Cromley (1995). Selv om vi i dag kan produsere klasseløse koropletkart adskillig lettere enn for 25 år siden, er dette en lite brukt løsning. Dette kan dels skyldes at kartografer aldri synes å ha blitt enige om hvorvidt det er ønskelig med klasseløse koropletkart. Dessuten har en vanligvis ikke mulighet til dette ved hjelp av kommersiell programvare da maksimalt mulig antall klasser normalt vil være mindre enn antallet ulike observasjonsverdier[4]. I Excel er det maksimale antall klasser 16. For de fleste datasett vil produksjon av klasseløse koropletkart dermed ikke være mulig i Excel.

Valg av klasseinndelingsmetode

Er det også fornuftig at standard metode for å bestemme klassegrenser er en kvantil inndeling? Evans (1977) påpeker at å inndele klasser i kvantiler synliggjør vanligvis en eller annen geografisk differensiering.

> Percentiles have been selected by cartographers wishing to play safe and make sure that some spatial differentiation was portrayed (Evans 1977:107).

Kartet i figur 1 er dannet ved å benytte en kvantil inndeling. Kartet fremhever regionale forskjeller med hensyn til andel unge av total befolkning tydelig. Det er en synlig regional differensiering for mellom kommuner i nord og kommuner i sør og mellom kyst- og innlandskommuner. Til sammenlikning fremstiller kartet i figur

---

[3] I MapInfo er dette standard oppsett ikke noe som er gitt en gang for alle, men noe som kan endres om bruker finner det hensiktsmessig.

[4] De fleste kommersielle programvarer har en øvre grense for mulige klasser (for eksempel er øvre grense 16 i MapInfo 4.1 og 64 i ArcView 3.1). Om klasseløse koropletkart skulle være mulig, burde det maksimale antall mulige klasser tilsvare antall ulike observasjonsverdier i brukers datasett.

2 de samme data men her har klassegrensene like intervaller. Dette kartbildet fremviser langt fra den samme regionale differensieringen og norske kommuner synes å ha stor likhet med hensyn til andel unge.

En ulempe ved Excels kvantilinndeling er at programmet ikke alltid lykkes med å gi et tilnærmet likt antall observasjoner i hver klasse. I figur 1 er det et sprang på 21 mellom minste og største antall observasjoner blant klassene. Dette kunne endres ved at bruker justerte klassegrensene, men dette er en umulig oppgave i Excel. Endring av klassegrenser kan gjøres i tegnforklaringen, men dette medfører ingen omgruppering av observasjonsenhetene (kommunene). Klassegrensene er alene bestemt ut fra programmets algoritmer for de to tilgjengelige klassifikasjonsmetodene og kan ikke endres.

Valg av grafisk fremstilling

Standard grafisk fremstilling er fornuftig i Excel. Når en ordnet variabel skal fremstilles må den symboliseres med en visuell variabel som er ordnet (Bertin, 1981). Gråtone (eller tetthet) er en ordnet visuell variabel – det er intuitivt for alle at mørke gråtoner representerer verdier større enn lyse gråtoner. Valg av grafisk fremstilling for standard kartet er derfor basert på en logikk om systematisk orden. I stedet for gråtoner kan en i Excel velge en gradert skala i andre farger, som for eksempel en skala fra lys til mørk rød. Brukeren har imidlertid ikke anledning til å velge farger for de ulike klassene (som en serie med grønn, brun, blå, –). Dermed forhindres brukeren i å 'bryte' det som de fleste oppfatter som 'kartografiske regler'. Farge er ikke en ordnet visuell variabel – det er ikke intuitivt at blått representerer en verdi større enn for eksempel grønn. Om bruker derimot endrer karttype til kvalitativt skravurkart kan de ulike klassene denne karttype genererer, symboliseres med ulike farger. Da representerer imidlertid ikke fargene en ordnet variabel men en kvalitativ variabel slik at fargene formidler kvalitativ likhet eller forskjell.

Det uheldige ved Excels standard symbolisering er at den laveste klassen symboliseres identisk med områder uten data. Om en ønsker å rette på dette ved for eksempel å gi den laveste klassen en svak gråtone, har en ingen mulighet til dette. Symboliseringen er gitt av programmet og kan ikke manipuleres med. En annen ulempe er at toveisskalaer som benyttes i bipolare koropletkart for å vise graderte skalaer over og under bestemte verdier ikke er mulig

Fleksibilitet

Ut fra betraktningene over kan vi konkludere med at koropletkart produksjon i Excel er basert på kartografiske 'regler'. Imidlertid er ikke Excel dermed funksjonell. Modulen synes å være implementert utfra spørsmålet: Hva ønsker bruker å utføre og hvordan kan programvaren tilby dette på en enklest mulig måte? Enkelheten eller brukervennligheten er oppnådd ved å tilby få funksjoner og et sett ferdige beslutninger som bruker ikke kan endre. Dette gjør Excel ufunksjonell for kartfremstilling. Baudouin (1987) fremholder blant annet at et kartografisk program bør være godt utrustet med hensyn til inndelingsmetoder. I Excels kartmodul tilbys kun metodene kvantil og like intervall inndeling. I følge Evans (1977) kan de fleste metoder for å bestemme klassegrenser betraktes som kompromisser av kvantil og like intervall inndelingen:

> . . . most class-interval systems (–) can be envisaged as compromise between percentile classes, with the desirable property of placing equal numbers of symbols in each class, and equal-interval classes, with the desirable property of equal width (Evans 1977:106)

Dette er likevel meget begrensende for en bruker som vil noe mer enn bare å produsere standard kart. Mange vil nok ønske å endre på klassegrenser om de for eksempel har behov for å finne den geografiske fordeling i henhold til bestemte verdier. For eksempel, om en bruker ønsker å lage et kart som viser hvilke kommuner som har færre enn 5000 innbyggere[5], har kartmodulen i Excel ingen mulighet til å imøtekomme et slikt ønske. Selv om brukergrensesnittet er enkelt kan det dermed ikke karakteriseres som funksjonelt om vi betrakter et brukergrensesnitt slik som Gould (1993):

> . . . [the user interface is] a conceptual link between human intention and what the computer can offer as a decision support environment (Gould, 1993:102).

---

[5] Referer debatten om sammenslåinger av norske kommuner med innbyggertall lavere enn 5000.

I stedet for at programvaren responderer på brukers intensjoner, fjernes brukeren som aktør i kartfremstillingsprosessen. Dette er i sterk motsetning til samtidens retninger innen kartografi om å *visualisere* geografiske data for å tilegne seg ny kunnskap i tillegg til effektivt å *kommunisere* et kjent budskap (se for eksempel MacEachren 1995, Kraak & Ormeling 1996). I den forbindelse er interaksjon mellom bruker og kart, som i disse tilfeller vil være en visning på skjerm, av meget stor betydning fordi aktøren og hennes persepsjon betraktes som betydningsfulle verktøy i prosessen å finne ny kunnskap.

> Human vision, instead of being considered a potential source of bias, has come to be recognized as a powerful tool for extracting patterns from chaos (MacEachren and Monmonier, 1992: 197).

Excel hindrer en bruker både med hensyn på å finne relevante geografiske mønstre og med hensyn på å presentere disse. Grunnen er snever funksjonalitet.

To kriterier for god klasseinndeling.

Prosessen å fremstille koropletkart er meget automatisert i Excel og mange brukere vil nok velge å benytte standard opsjonene. Dermed oppstår situasjonen beskrevet av Kennedy:

> Default classifications are provided with automated cartography programs by many software developers to 'help' the first-time user with the software. Unfortunately, many users accept the default classification as the 'correct' classification (Kennedy 1994:20).

Hva er imidlertid en korrekt klassifikasjon? Et objektiv kriterium på om klassifikasjonen er god er å beregne klasseinndelingens statistiske tilpasning i forhold til de opprinnelige data. Desto større homogenitet innen klassene (liten varians) og desto større heterogenitet mellom klassene (stor varians), desto bedre klassifikasjon etter disse kriterier. GVF, eller *Goodness of Variance Fit*, er et slikt mål.

GVF verdien beregnes på følgende måte (Dent 1996:136):

1. Beregn aritmetisk gjennomsnitt ($\bar{x}$) for hele datasettet og beregn summen av de kvadrerte avvik mellom observasjonsverdiene (x_i) og gjennomsnittet:

$$\sum (x_i - \bar{x})^2$$

Denne verdi kalles SDAM (squared deviations, array mean).

2. Beregn aritmetisk gjennomsnitt for hver klasse (Z_c). For hver klasse beregnes summen av de kvadrerte avvik mellom observasjonsverdiene (x_i) og klassens aritmetiske gjennomsnitt ($x_i - Z_c$). Til slutt summeres det for alle klassene:

$$\sum \sum (x_i - Z_c)^2$$

Denne verdi kalles SDCM (squared deviation, class means).

3. Beregn *the goodness of variance fit* (GVF):

$$GVF = \frac{SDAM - SDCM}{SDAM}$$

Maksimalverdien for GVF er 1,0 og de beste klassifikasjonsløsninger vil ha verdier opp mot 1,0. GVF kan dermed benyttes for å optimalisere klassifikasjonen (Jenks & Caspall 1971, Jenks 1977) eller for å sammenlikne nøyaktighet mellom ulike klassifikasjoner (Smith 1986, Declercq 1995). Om vi beregner GVF verdier for datasettet for årstallet 1997 inndelt i fem klasser gir kvantil inndelingen en GVF verdi på 0.82, mens like intervaller gir en GVF verdi på 0.62.

Hvor stor bør imidlertid GVF verdien være før vi kan konkludere med at klassifikasjonen er god? Declercq (1995:922) konkluderte at en GVF verdi på 0,95 kan betraktes som normativt for en god klassifikasjon. Generelt øker GVF verdien med antall klasser[6], derfor behøvde Declercq minst seks klasser for å oppnå en godhet på 0,95 for et datasett med 308 observasjoner. Når datasettet 'Andel unge (0 til 20 år) i 1997 i prosent av total befolkning' som består av 435 observasjoner grupperes i fem klasser er en godhet på 0,95 følgelig ikke oppnåelig. Ved å benytte Jenks optimale klassifikasjon ble det oppnådd en GVF verdi på 0,916 som dermed er GVFs maksimalverdien for dette datasett inndelt i fem klasser[7]. Dette er dermed den klassifi-

[6] GVF = 1 når antall klasser er lik antall observasjoner.

[7] Dette er utført på en implementert Visual Basic kode basert på Jenks (1977) og Lindberg (1990).

kasjon som best gjengir den iboende statistiske fordeling til variabelen. Om vi sammenlikner GVF verdiene oppnådd ved de to tilgjengelige klassifikasjonsmetodene i Excel, ser vi at GVF verdien kalkulert fra kvantil inndelingen er adskillig nærmere maksimalverdien enn GVF verdien kalkulert fra like intervall metoden. Dermed har vi nok et argument som støtter bruken av kvantil inndelingen som standard for dette datasett.

Dette kriteriet for god klassifikasjon kan sies å tilhøre det som benevnes som kommunikasjonsparadigmet innen kartografi. Innen dette paradigme er målet med kartografi å kommunisere et bestemt budskap effektivt. Dette forutsetter at budskapet er kjent på forhånd og at det finnes ett optimalt kart for dette. Kartografenes mål er innen kommunikasjonsparadigmet å komme frem til det optimale kart. GVF verdien fungerer som en angivelse om klasseinndelingens godhet og burde derfor bli presentert i en tilbakemelding som vurderer de valg bruker har gjort. Dette er imidlertid fraværende i Excel og i alle andre kartografiske programpakker denne forfatter kjenner.

Et annet, subjektivt kriterium for god klassifikasjon er om kartet får frem den geografiske fordeling brukeren ser etter. Dette kriteriet for god klassifikasjon kan sies å tilhøre det som benevnes som visualiseringsparadigmet innen kartografi. Innen dette paradigmet er ikke målsettingen å formidle et budskap mest mulig effektivt. Budskapet er, som poengtert av MacEachren og Ganter, ukjent:

> For cartographic visualization the message is unknown and, therefore, *there is no optimal map!* The goal is to assist an analyst in *discovering patterns and relationships* in the data (MacEachren & Ganter 1990: 65).

Innen visualiseringsparadigmet er målsettingen å perspektivisere datasettet ved hjelp av ulike kartografiske fremstillinger. Resultatet av denne prosessen blir da forhåpentligvis at brukeren oppdager mønstre og relasjoner i datasettet og dermed oppdage ny kunnskap for det fenomen som studeres.

Valg av antall klasser, klassegrenser og grafisk fremstilling påvirker det geografiske mønsteret som dannes i kartet. Geografiske forskjeller som finnes i datasettet kan ubevisst eller bevisst skjules eller fremheves ved endringer av disse tre parametre. Når en i Excel, på grunn av manglende fleksibilitet, nærmest oppfordres til å benytte kun standard kartet er dette nærmest uetisk noe Monmonier sterkt tar avstand fra:

> . . . the decision to present a **single** cartographic viewpoint can be a decision fraught with important ethical overtones (Monmonier, 1991).

Sandberg (1998) har satt fokus på dette i Norge ved å fremlegge sitt hovedpoeng: Kartografer gjør etiske valg. Et kart er ikke et speilbilde av virkeligheten, men en representasjon av kartografens tolkninger eller oppfatninger av denne. Om kartproduktet skal presenteres for et større publikum er dermed kartografen med på å forme andres virkelighetsoppfatninger.

Konkluderende perspektiv

Skal en være i stand til å oppdage geografiske sammenhenger eller mønstre fra et eller flere datasett, må en kunne ha stor grad av interaktivitet mellom programvare og bruker. Brukervennligheten til en slik programvare bør derfor etterstrebes ved å legge til interaktivitet heller enn å redusere denne ved at valgmuligheter og fleksibilitet fjernes. Enkelhet kan likevel oppnås ved å implementere et brukergrensesnitt der designprosessen er dekomponert til noen få relevante og logiske trinn, gjerne i et veiviser miljø som bruker kan følge. Med hensyn til fremstilling av koropletkart, er det tre relevante, logiske trinn en slik veiviser bør bestå av: valg av antall klasser, valg av metode for å bestemme klassegrenser og valg av grafisk fremstilling. Om dette veiviser miljøet har som Excel fornuftige standard valg, men i tillegg hjelpefunksjoner og tilbakemeldinger tilgjengelig for brukeren slik at hun eventuelt kan reversere prosessen, vil brukers forståelse av kartfremstillingsprosessen sannsynligvis øke. Et slikt brukergrensesnitt har pedagogisk potensiale som påpekt blant annet av Raper:

> Following such a path and gaining experience with the alternative options is an excellent way to improve a user's end-to-end understanding of the components of spatial data processing. (Raper 1991: 111)

Om fremtidens kartografiske programpakker inneholder et slikt pedagogisk brukergrensesnitt har vi ikke noe å frykte ved en allminneliggjøring av GIS eller kartografisk programvare. Utfallet av at hvem som helst er i stand til å lage kart kan da bli større allmenn kunnskap om god kartdesign og en bevissthet om kartets rolle som visualiseringsverktøy for å oppnå ny kunnskap eller som kommunikasjons-medium for å formidle kunnskap.

Referanser

Baudouin, A. (1987) Klasseinndeling i Kartografi. *Kart og Plan* Vol. 47, No. 3, pp. 321–326.

Bertin, J. (1981) *Graphics and Graphic Information Processing.* Walter de Gruyter.

Cromley, R. G. (1995) Classed versus Unclassed Choropleth Maps: A Question of How many Classes. *Cartographica.* Vol. 32, No. 4, pp. 15–27.

Declercq, F. A. N. (1995) Choropleth Map Accuracy and the Number of Class Intervals. *17th International Cartographic Conference «Cartography Crossing Borders».* Barcelona, September 1995.

Dent, B. D. (1996) *Cartography – Thematic Map Design.* Wm. C. Brown Publishers.

Dorling, D. & Fairbairn, D. (1997) *Mapping: Ways of Representing the World.* Longman.

Evans, I. S. (1977) The selection of class intervals. *Transactions of the Institute of British Geographers.* Vol 2, pp. 98–124.

Gould, M. D. (1993) Two Views of The User Interface. In *Human Factors in Geographical Information Systems.* Eds. Medyckyj-Scott, D. & Hearnshaw H. M, pp. 101–110. London: Belhaven Press.

Jenks, G. F. (1977) Optimal Data Classification for Choropleth Maps. *Occational Paper No 2, Department of Geography University of Kansas Lawrence, Kansas.*

Jenks, G. F. & Caspall C. (1971) Error on Choroplethic Maps: Definition, Measurement, Reduction. *Annals of the Association of American Geographers.* Vol. 61, No. 2, pp. 217–244.

Kennedy, S. (1994) Unclassed Choropleth Maps Revisited / Some guidelines for the construction of unclassed and classed choropleth maps. *Cartographica.* Vol. 31, No. 1, pp. 16–25.

Kraak, M. J. & Ormeling, F. J. (1996) *Cartography – Visualization of spatial data.* . Longmann.

Lang, L. (1995) The Democratization of GIS – Bringing Mapping to the Masses. *GIS World.* Vol. 8, No. 4, pp. 62–67.

Lindberg, M. B. (1990) Fisher: A Turbo Pascal Unit for Optimal Partitions. *Computers & Geosciences.* Vol. 16, No. 5, pp. 717–732.

MacEachren, A. M. (1995) *How Maps Work – Representation, Visualization and Design.* The Guilford Press.

MacEachren, A. M. & Ganter, J. H. (1990) A Pattern Identification Approach to Cartographic Visualization. *Cartographica.* Vol. 27, No. 2, pp. 64 – 81.

MacEachren, A. M. & Monmonier, M. (1992) Introduction. In: *Geographic vizualization.* Bethesda, Md.: American Congress on Surveying and Mapping. In serie: *Cartography and Geographical Information Systems.* Vol. 19, No. 4, pp. 197–200.

Mersey, J. E. (1990) *Colour and Thematic Map Design – The Role of Colour Scheme and Map Complexity in Choropleth Map Communication.* Cartographica. Vol. 27, No. 3. Monograph 41.

Monmonier, M. (1991) Ethics and Map Design: Six Strategies for Confronting the Traditional One-Map Solution. *Cartographic Perspectives.* No. 10, pp. 3–8.

Monmonier, M. (1995) *Drawing the line: tales of maps and cartocontroversy.* New York. Henry Holt and Company, Inc.

Raper, J. F. (1991) User interfaces. In *Handling Geographical Information*, ed. Massey, I. & Blakemore, M. pp. 102–114. Longman.

Sandberg, P. (1998) Etikk i kartbransjen. *Foredrag på Kartdagene 1998.*

Smith, R. M. (1986) Comparing Traditional Methods for Selecting Class Intervals on Choropleth Maps. *Professional Geographer*, Vol. 38, No. 1. pp. 62–67.

Tobler, W. (1973) Choropleth maps without class intervalls. *Geographical Analysis.* Vol 5, No 3, pp. 262–265.

Assessing the functionality of choropleth mapping using Excel

Jan Ketil Rød

Abstract

For several years users of standard software for spreadsheets have been able to visualise their tabular data as maps due to a new tool: statistical mapping. One such standard software that has imbedded mapping capabilities is Excel. An evaluation of Excel's functionality for statistical mapping is undertaken. Functionality is herein taken to mean the software's ability to respond to the user's needs and aims. The evaluation deals with choropleth mapping, a mapping process, which is broken down into three decisions: the choice of number of classes, the choice of method for class breaks determination, and choice of symbolisation. How the Excel mapping environment performs each of these steps is investigated in comparison with cartographic theory. Finally, some reflections on the influence the democratisation of mapping might have on map design are put forward.

Key words: Choropleth mapping, map design, visualisation, democratisation.

Introduction

The point of departure for this article is the public availability of statistical data and the ability for everyone to map these data. The latter is made possible as common software packages for spreadsheets and statistics have a mapping module embedded. The users of for instance the mapping package MS Excel are themselves now able to make choropleth maps, dot density maps, proportional point symbol maps, chorochromatic maps, bar or pie charts maps. These types of maps might all be described as "statistical maps": maps who's purpose is to represent the *geographical* distribution a data set has: e.g. volume of sales, demographic data, etc. Because MS Excel and similar software packages are more or less standard office equipment, a very different situation occurs today: anyone amongst the 50 million users of spreadsheet packages, can now individually produce statistical maps. This situation was commented on in *The Democratization of GIS – Bringing Mapping to the Masses* (Lang, 1995). The democratisation, or making mapping common, might result in the production of poor maps. Another possible outcome is the diffusion of cartographic knowledge to the general public. According to Monmonier (1995:1) maps are less subject to critical examination than written forms of communication. When everybody is able to produce maps, map-making might become demystified (Dorling and Fairbairn 1997:165) in such a way that the idea of the map as a reflection of reality is replaced with the view of mapping as a statement just like any other statement expressed textually or verbally. Thereafter an increase in the number of critical voices is likely to occur.

This article is restricted to the assessment of the functionality of one common software package, which allows for statistical mapping: MS Excel. Functionality, here, is taken to mean the user's flexibility related to the choices that have to be made and the way the system responds to these. Are the users able to carry out all their intentions; do they get any feedback regarding the consequences on their choices? The main subject in this article is choropleth mapping. Choropleth maps for time series are not discussed. How the varying size of the administrative units may influence the visual perception of the map is also not taken up here.

Assessing choropleth mapping in Excel

When choropleth mapping is the subject for evaluation it is adequate to breakdown the process, which can be done by separating the three main decisions comprising it (Baudouin 1987):

1. The selection of number of classes
2. The selection of classification method
3. The selection of symbolisation

How are these decisions grounded and under which conditions are they taken for a user of the Excel package and what is the result? First, the question will be evaluated regarding the default map, which the Excel mapping module offers: a map designed exclusively on decisions taken by the system. Thereafter, the issue will be evaluated in accordance with the user's flexibility to redo these three decisions. For the purpose of this study Excel 97 is used.

Assessing the default map design in Excel and the user's flexibility

Figure 1 shows an example on the Excel default map (a choropleth map according to the quintile classification) from the variable 'the percentage of young people (0 to 20 years) of the total population in 1997' for Norwegian municipalities[1]. In order to produce this map, the columns in the spreadsheet holding the variable and the municipalities' identification number must be marked. By clicking on the mapping icon (represented by a globe) a map is generated. As a default representation, a choropleth map with five classes is made. The class intervals are determined according to the principle of an equal or near equal number of observations in each class, a classification method usually called *percentile* or *quantile*[2]. The symbolisation is graded by shades of grey from white through dark grey. The number of observations falling into each class is shown in brackets in the legend.

Cartographic praxis and research have resulted in a more or less consensus regarding the "rules" for good map design. Is the default map produced by Excel in accordance with these "cartographic rules"?

Selecting number of classes

Since it is only possible to discriminate between a limited number of shades of grey, perception states the reason for limiting the number of classes. As a compromise between different results and theories five classes are found suitable while the number of classes might be slightly higher if a colour scheme is used (Mersey 1990). The default number of classes is five in the Excel mapping module and the default symbolisation is an ordered series in shades of grey. The "mother" software package MapInfo uses seven classes as default and an ordered colour scheme from light to dark red[3]. Evans (1977) argues that increasing the number of classes increases map complexity and consequently the number of classes should also be a question on how familiar the user is with "reading" graphical information.

> Within the range four to ten classes, a decision should be influenced by the intended audience, the technical means available, and the spatial pattern of the distribution. A simple clear-cut map with four or five classes may be better for an unsophisticated audience, inexperienced at reading graphics. Trained eyes may appreciate the extra information which seven or eight classes portray (Evans 1977:100).

---

[1] The maps in this article represent variables generated from demographic variables available from the municipality database ("kommunedatabasen"), which is gathered by the Norwegian Social Science Data Service ("Norsk Samfunnsvitenskapelig Datatjeneste – NSD"). The coordinates for the municipality borders are also provided by the NSD. The basis maps are generated in MapInfo, the Excel mapping modules "mother" software package.

[2] "Percentiles" is a term often used to avoid confusing 'quartiles' (four classes) and 'quintiles' (five classes).

[3] In the MapInfo package the default number of classes is not finally determined as the user might alter it.

It is likely that users of the MapInfo software are more sophisticated than ordinary users of Excel, who are also more likely to be inexperienced in cartography. It is therefore reasonable that the default number of classes is less in the Excel mapping module than in the “mother” software MapInfo.

With regard to the selection of number of classes, Tobler (1973) suggested a method for classless choropleth map, initiating a debate recently revisited by Kennedy (1994) and Cromley (1995). Although we are increasingly able to produce classless choropleth maps with considerable ease today compared to 25 years ago, classless choropleth maps are seldom used. This might be a result of cartographers never accomplishing an agreement on whether or not classless choropleth map is desirable. Furthermore, it is not possible to create classless choropleth maps by using available commercial mapping packages since the maximum number of classes will often be less than the number of unique observation values[4]. In Excel, as the maximum number of classes is 16, the production of classless choropleth map will not be possible for any dataset having an exceeding number of unique observation values.

Selecting classification method

Is it reasonable that the default classification method is the quantile method? Evans (1977) points out that grouping the observations into quantiles usually make visible some spatial differentiation.

> Percentiles have been selected by cartographers wishing to play safe and make sure that some spatial differentiation was portrayed (Evans 1977:107).

The map in Figure 1 is created according to a quantile division. The map emphasizes regional differentiations in the percentage of young people of the total. There is a clear regional differentiation between municipalities in the north and south and between coastal and inland municipalities. In comparison, the map shown in Figure 2 is based on the same data but the classes are delimited according to equal intervals. As a result, the map shows the regional differentiation shown in Figure 1 to a lesser extent, but represents the Norwegian municipalities as being rather similar with regard to the percentage of young people.

---

[4] Most commercial mapping packages have an upper limit for the number of classes (i.e. 16 in MapInfo 4.1 and 64 in ArcView 3.1). To be able to produce classless choropleth map, the software’s upper limit for number of classes must be high enough to include all unique observation values in the dataset.

A disadvantage in the way the Excel mapping module groups observations according to the quantile method is that the method does not always succeed in dividing the observations equally among the classes. In Figure 1 there is a difference of 21 units between the lowest and highest number of observations among the classes. The users would have been able to alter the number of observation falling into each class by adjusting the class breaks, but this is not feasible in Excel. The class breaks in the legend, may be altered but this does not alter the groupings of units (municipalities). Class breaks is solely set by the software package's algorithms and is thus out of the user's control.

Selecting symbolisation

The default symbolisation is reasonable in Excel. An ordered variable must be represented, according to Bertin, by an ordered visual variable (Bertin, 1981). Shades of grey (or density) is an ordered visual variable – and it is generally known by intuition that darker shades of grey represent higher values than lighter shades of grey. The choice of symbolisation for the default map is thus based on logic of systematically order. Instead of shades of grey, a graded shading within a colour may be selected, like, e.g. a scheme from bright to dark red. However, users cannot alter the ordered sequence from light to dark by selecting particular colours for the different classes (like a scheme with green, brown, blue). Consequently, users are prevented from "violating" what is mostly understood as "cartographic rules". Colour is not an ordered visual variable; one does not sense intuitively that the colour blue represents a higher value than, for instance, the colour green. On the other hand, if users represent the variable with a chorochromatic map the various classes this type of representation generates can be symbolised using different colours. In a choropleth map, however, the colours do not represent an ordered variable but a qualitative variable so that the colours transcribe qualitative equality or difference.

Unfortunately, the lowest class is symbolised identical with areas with no data by the default symbolisation offered by Excel. If one wants to correct this by for instance giving the lowest class a bright shade of grey leaving white for the "no data" class there is no opportunity to do so. The system sets the symbolisation and it cannot be altered. Another drawback is that bipolar symbolisation schemes, which might be used to indicate units respectively above and below certain values, are also not possible to produce.

Flexibility

Based on the considerations above, it is evident that the choropleth map production is based on cartographic “rules”. This fact, however, does not make Excel functional. The mapping module seems to be implemented to fulfil the issue: What do users need to carry out and how can the software accomplish these needs as simply as possible? Offering a limited number of options and setting ready-made decisions which the user cannot alter**,** is overly simplistic but makes Excel worthless as a tool for mapping. Baudouin (1987) points out that a cartographic package ought to be well equipped regarding classification methods. The Excel mapping module only offers the methods quantile and equal interval. According to Evans (1977) most of the methods for class breaks determination can be regarded as compromises of the quantile and equal intervals methods:

> ... most class-interval systems (…) can be envisaged as compromise between percentile classes, with the desirable property of placing equal numbers of symbols in each class, and equal-interval classes, with the desirable property of equal width (Evans 1977:106).

Still, this is indeed limiting for a user wanting something more than to merely produce default maps. Many users would probably like to be able to alter the class breaks in order to visualise the geographical distribution in accordance with certain values. For instance, if a user needs a map showing the municipalities having less than 5000 inhabitants[5] the mapping module in Excel cannot accommodate such a request. Even though the user interface is simple, it cannot be characterised as functional if we regard the user interface in the same manner as Gould (1993) does:

> ... [the user interface is] a conceptual link between human intention and what the computer can offer as a decision support environment (Gould, 1993:102).

Instead of responding to the user’s intensions, the user is a non-active figure in the mapping process, which is in stark contrast to contemporary directions within cartography on *visualising* geographical data in order to achieve new knowledge in addition to effectively *communicate* a known message (see for example MacEachren 1995, Kraak and Ormeling 1996). As the users perception is regarded as essential, the interaction between user and map, which in these instances will be map displays on screen, will in this context mean a great deal in terms of acquiring new knowledge.

> Human vision, instead of being considered a potential source of bias, has come to be recognized as a powerful tool for extracting patterns from chaos (MacEachren and Monmonier, 1992: 197).

Due to limited functionality, users are hindered both in exploring the data to find relevant geographical patterns, and in presenting these results.

Two criteria for good classification

The process of making choropleth maps is highly automated in Excel and many users will probably select the default options, which consequently bring the situation described by Kennedy into being:

> Default classifications are provided with automated cartography programs by many software developers to 'help' the first-time user with the software. Unfortunately, many users accept the default classification as the 'correct' classification (Kennedy 1994:20).

Nevertheless: what is, however, a correct classification? An objective criterion whether the classification is good or not, is to calculate how well the groupings fit the original data. The more homogeneity within classes (little variation) and the greater the heterogeneity between classes (huge variation), the better the classification according to these criteria. GVF, or *Goodness of Variance Fit*, is a measure of classification 'fitness'.

The GVF value is calculated as follow (Dent 1996:136):

1. Calculate the arithmetic mean ($\bar{x}$) for the variable and calculate the sum of the squared deviations between observation values (x_i) and the mean:

   $$(x_i \; 4 \; \bar{X})^2$$

   This value is called SDAM (squared deviations, array mean).

2. Calculate the arithmetic mean within each class ($\bar{Z}_c$). For each class calculate the sum of the squared deviations between observation values (x_i) and the class' arithmetic mean ($x_i \; 4 \; \bar{Z}_c$). Finally, the sum of all classes:

   $$(x_i \; 4 \; \bar{Z}_c)^2$$

   This value is called SDCM (squared deviation, class means).

---

[5] Refer to the debate on merging Norwegian municipalities having less than 5000 inhabitants.

3. Calculate the goodness of variance fit (GVF):

$$GVF = \frac{SDAM - SDCM}{SDAM}$$

The maximum value for GVF is 1,0 and the best solutions on data classification will result in GVF values approaching 1,0. Consequently, GVF might be applied to optimise the classification (Jenks & Caspall 1971, Jenks 1977) or to compare the accuracy between different classifications (Smith 1986, Declercq 1995). If we calculate GVF values for the variable *percentage of young people (0 to 20 years) of the total population in 1997* grouped into five classes, the quantile division results in a GVF value at 0.82 while the equal interval division results in a GVF value at 0.62.

How high ought the GVF value be for a classification to be considered as good? Declercq (1995:922) concluded that a GVF value at 0,95 could be taken as a norm for a good classification. Generally, the GVF value increases with an increasing number of classes[6], and consequently, Declercq needed six classes or more to obtain a GVF value better than 0,95 on a variable with 308 observations. By using an optimal classification method (Jenks 1977, Lindberg 1990) a GVF value at 0,916 was obtained for the classification of the variable *percentage of young people (0 to 20 years) of total population in 1997* grouped into five classes[7]. The classification, with which this GVF value is achieved, is thus the classification that best renders the inherent statistical distribution for this variable. However, a GVF value better than 0,95 is not achievable for this variable, consisting of 435 observations, when grouped into five classes. If we compare the GVF values achieved by the two classification methods available in Excel, it is noticeable that the GVF value resulting from the quantile division is closer to the possible maximum value than the GVF value resulting from the equal interval method. Consequently, we have an argument supporting the choice of classification method.

This criterion for good classification might be regarded as belonging to what is called the communication paradigm within cartography. Within this paradigm, the aim of cartography is to communicate a certain message effectively, which presupposes an a priori known message and that there exist an optimal map for this message. The cartographers' objective is within

---

[6] GVF = 1 when the number of classes equals the number of observations.

[7] This is carried out using an implemented Visual Basic code based on Jenks (1977) and Lindberg (1990).

the communication paradigm to succeed in finding the optimal map. GVF value is functioning as a specification on the classification's goodness and should therefore be included, as an assessment on the class interval selections done, in a response to the user. However, this is an absent function in Excel as all the other mapping packages known to this author.

Another, subjective criterion for good classification is whether the map shows the geographical pattern the user is looking for. This criterion for good classification may belong to what is called the visualisation paradigm within cartography. Within this perspective, the aim is not to communicate a message as effectively as possible. The message is noted by MacEachren and Ganter, unknown:

> For cartographic visualization the message is unknown and, therefore, *there is no optimal map!* The goal is to assist an analyst in *discovering patterns and relationships* in the data (MacEachren & Ganter 1990: 65).

Within the visualisation paradigm, the aim is to put the data into different perspectives by means of several cartographic portrayals. The result of this process is hopefully that the user discovers patterns and relations in the data and thereby discovers new knowledge about the subject matter.

The selection of number of classes, class breaks and symbolisation affect the resulting geographical pattern, which appears in the map. Geographical differences which are present in the dataset can unconsciously or consciously be concealed or enhanced by altering these three parameters. When an Excel user, due to the reduced flexibility of the software, is practically forced to make use of the default map, this may be regarded as almost unethical: Monmonier reiterates these claims:

> ... the decision to present a **single** cartographic viewpoint can be a decision fraught with important ethical overtones (Monmonier, 1991).

Sandberg (1998) has put this issue into focus in Norway by presenting his main point: Cartographers take ethical choices. A map is not a mirror of reality, but a representation of the cartographer's interpretations or comprehensions of it. If the map is to be presented for a wider public, the cartographer contributes in shaping others understanding of the real.

Concluding perspective

In order to discover geographical relations or patterns based on one or several datasets, it is necessary to maintain a high level of interactivity between the software and user. Such software packages ought to be developed aimed at providing substantially greater user-friendliness. Key features of such a product should include adding interactivity rather than reducing it by removing choice options and flexibility. Simplicity can still be obtained by implementing a user-interface where the design process is broken down to a few, but relevant and logical steps, readily available in a “wizard” environment making it easy for the user to follow. With regard to choropleth mapping, there are three relevant, logical steps such a wizard function should include: the selection of the number of classes, class intervals selection and the selection of symbolisation scheme. If such a wizard environment has, as Excel, reasonable standard options, but also ˈ and available for the userˈ a help function with responses allowing the user to reverse the process if necessary, the user’s understanding of the mapping process will in all likelihood be increased. Such a user interface does have educational potential as pointed to by Raper:

> Following such a path and gaining experience with the alternative options is an excellent way to improve a user’s end-to-end understanding of the components of spatial data processing (Raper 1991: 111).

If the prospective mapping packages have such an educational user interface embedded, there is nothing to be afraid of by the democratisation of GIS or mapping packages. The result of anybody being able to produce maps may well result in an increased general knowledge about the nature of good map design and an awareness of how maps function as visualisation tools for gaining new knowledge and as communication devises that pass on knowledge.

Critical comments

> If we calculate GVF values for the variable *percentage of young people (0 to 20 years) of the total population in 1997* grouped into five classes, the quantile division results in a GVF value at 0.82 while the equal interval division results in a GVF value at 0.62 (Rød, 1999: 40).

The calculation of GVF values as referred to in the citation above, was done using an implemented visual basic code witch later, with modifications, have been included in the GIB mapping package. Unfortunately, after the publication of this article, I discovered an error in the implemented code regarding the calculation of GVF values. In Table 1 below, the first column shows the values as they appear in Rød (1999) and the second column shows corrected GVF values for identical class breaks as in Rød (1999). The quantile division performed by the mapping module in MS Excel succeed, as commented (Rød, 1999: 39), very poorly in assigning an equal number of observations per classes (see Figure 1 in Rød, 1999). As the methods for the classification methods implemented in GIB result in other groupings (especially for the quantile method), the calculated GVF values for these are found in the third column in Table 1.

	Erroneous values (from Rød, 1999)	Corrected values	Results from the grouping done by GIB
Quantile	0,82	0,862	0,874
Equal Interval	0,62	0,877	0,880
Optimal	0,916	0,917	0,917

Table 1: Values for GVF for the variable *Number of inhabitants aged 0 to 20 years in Norway in 1997* grouped into five classes.

The following statement:

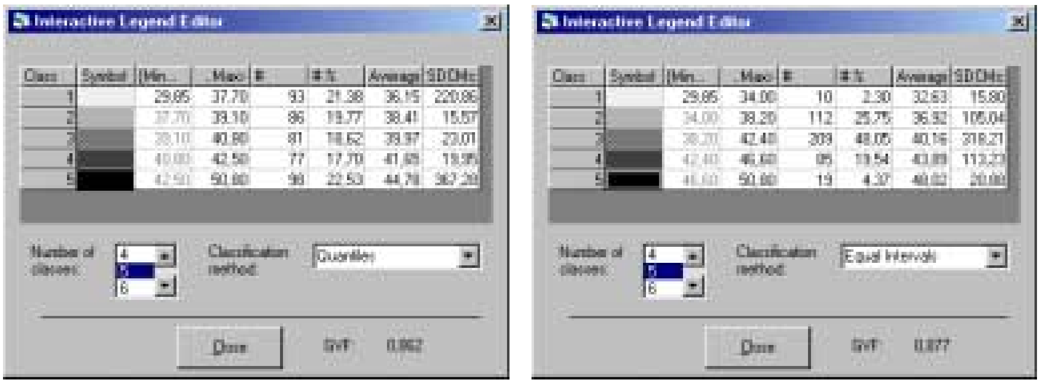
> If we compare the GVF values achieved by the two classification methods available in Excel, it is noticeable that the GVF value resulting from the quantile division is closer to the possible maximum value than the GVF value resulting from the equal interval method. Consequently, we have an argument supporting the choice of classification method (Rød, 1999: 41).

is consequently refuted. There is hardly any difference between the classification methods regarding how they score on GVF values for this data set. However, the following statement:

> The map in Figure 1 is created according to a quantile division. The map emphasizes regional differentiations in the percentage of young people of the total. There is a clear regional differentiation between municipalities in the north and south and between coastal and inland municipalities. In comparison, the map shown in

> Figure 2 is based on the same data but the classes are delimited according to equal intervals. As a result, the map shows the regional differentiation shown in Figure 1 to a lesser extent, but represents the Norwegian municipalities as being rather similar with regard to the percentage of young people (Rød, 1999: 38-39).

is still true. The idea to classify a variable to enhance differentiations, however, is a practice I would not recommend without carefully controlling that the geographical differentiation created resembles the "real" situation. The more accurate the classification is, the more likely is it that the map resembles the actual geographical pattern. As both classification method here are nearly equal and below the recognised threshold value for a good classification (i.e. 0,95), it is likely that both method, although in different ways, gives a biased picture on the geographical situation. Figure II.1 below shows how.



Class	Symbol	[Min...	...Max]	#	# %	Average	SDCMc
1	□	29.85	37.70	93	21.38	36.15	220.86
2	░	37.70	39.10	86	19.77	38.41	15.57
3	▒	39.10	40.80	81	18.62	39.97	23.01
4	▓	40.80	42.50	77	17.70	41.65	19.95
5	█	42.50	50.80	98	22.53	44.78	367.20

Number of classes: 5

Classification method: Quantiles

GVF 0.862

Class	Symbol	[Min...	...Max]	#	# %	Average	SDCMc
1	□	29.85	34.00	10	2.30	32.63	15.80
2	░	34.00	38.20	112	25.75	36.32	105.04
3	▒	38.20	42.40	209	48.05	40.16	318.21
4	▓	42.40	46.60	85	19.54	43.89	113.22
5	█	46.60	50.80	19	4.37	48.02	20.00

Number of classes: 5

Classification method: Equal Intervals

GVF 0.877

Figure II.1: The way the quantile method (left) and equal interval method (right) generate classification error. While the quantile division results in SDCMc values highest for the extreme classes, the equal interval division results in SDCMc values highest for the mid classes. High SDCMc values indicate concealed variation.

As seen from the SDCMc values, the quantile method emphasise a difference between geographical units in the lower class and the upper class and consequently, the SDCMc values for the lower and upper class are much higher than the rest. In comparison, the error contribution for the equal interval classification is highest for the mid-class but also high for the second and the fourth class, which is the reason why the "equal interval" map shows to a less extent the regional differentiation shown in the "quantile" map, but represents the Norwegian municipalities as being rather similar with regard to the percentage of young people.

Article III

Preface

Rød, J.K. 1998. Le troisième choix ... *Papers from The Department of Geography University of Trondheim – New series B, No. 43.* (Corrected edition, originally published in *Comité Français de Cartographie.* 156: 98-102).

The third choice …

Preface

I prepared the paper *Le Troisième choix* ... for the *Colloque 30 ans de sémiologie graphique, Paris, 12-13 Désembre 1997*. As I was told the contributions were going to be published at *CyberGeo*, which is a bilingual electronic journal, I prepared and submitted both an English and a French version. However, only the English version was published at *CyberGeo*. The French version was published in *Bulletin du Comité Français de Cartographie* 156: 98-102, unfortunately, not without serious errors. The English and French versions were not identical, neither in text, nor in Figures and Tables included. Unfortunately, *Bulletin du Comité Français de Cartographie* made a total mess out of the Figures and Tables resulting in using figure and table captions, which did not match the figures and tables printed, using Figures from the English version also in the French, omitting a table in the French, and ruin the order of succession of Figures and Tables.

A correct version of the paper was thus published in *Papers from The Department of Geography University of Trondheim New series B* No. 43 and is the paper included in this thesis. The English version is a translation of this correct version of the French paper, and not the English version published in CyberGeo.

CyberGeo is found on the url: http://www.cybergeo.presse.fr/revgeo2.htm

ARBEIDER FRA GEOGRAFISK INSTITUTT, UNIVERSITETET I TRONDHEIM

PAPERS FROM THE DEPARTMENT OF GEOGRAPHY, UNIVERSITY OF TRONDHEIM

Ny serie B - New series B

No. 43

JAN KETIL RØD

THE THIRD CHOICE...

LE TROISIEME CHOIX...

Abstract

Rød, J.K. 1998. The third choice... In: Colloque 30 ans de sémiologie graphique, Paris, 12-13 Décembre 1997.
http:/www.cybergeo.presse.fr/revgeo/semiogra/semiogrf.htm
(ISSN 0809-2966) [Language: English]

30 years ago Jacques Bertin launched the theory of graphic semiology. This theory has formed one of the pillars in cartographic education and in practical cartographic work. The objective of this paper is not to review this fruitful adoption but rather to draw attention to an inconsistency betweenBertin's term *organising level* and the corresponding concept from the literature of statistics, *measurement system.* The fact that the former is three-levelled and the latter is four-levelled obscures the schemata used for cartographic symbolisation. This issue becomes more urgent in constructing a rule-base for an expert system accommodated for statistical mapping. A framework for understanding the issue is provided. It reviews some of the obvious inconsistencies and investigates a possible solution, that is a symbolisation schema that might be applied as a cartographic rule base. This suggests a symbolisation schema applied to two main purposes of maps: those used as a visualisation tool and those used as a communication device.

Abstrait

Rød, J.K. 1998. Le troisième choix... Corrected version of paper published in: *Bulletin du Comité Français de Cartographie* 156, 98-102.
(ISSN 0809-2966) [Language: French]

Il y a 30 ans Jacques Bertin publiait Sémiologie Graphique. Cette théorie s'est imposée comme primordiale, tant dans le domaine de l'enseignement de la cartographie, que dans celui de la cartographie appliquée. Cet article ne reviendra pas sur la fructueuse utilisation de la Sémiologie. Nous entendons montrer la contradiction entre les niveaux d'organisations proposés par Bertin, et ceux utilisés par les staticiens et les cartographes anglo-saxons. En effet, si Bertin identifie trois niveaux d'organisation les cartographes anglo-saxons en utilisent quatre. Dans une optique d'élaboration de règles de base pour un système expert destiné à l'élaboration de cartes thématiques. Ce floue dans la définition des niveaux d'organisation est un problème fondamental. Nous analyserons les principales contradictions entre les deux systèmes de mesures et proposerons des solutions. Cette réflexion aboutit à un schéma de synthèse utilisable comme règle de base. Ce schéma s'applique aux deux fonctions de la carte : outil de visualisation et moyen de communication.

Introduction

Cette présentation est une synthèse de l'article écrit à l'occasion du trentième anniversaire de la Sémiologie Graphique. L'article traite des règles de symbolisation applicables pour faire les cartes statistiques. Comme Baudouin (1987) l'a affirmé, les décisions qui doivent être prises pour la réalisation de telles cartes, peuvent être groupées sous la forme de trois questions :

- Le choix du nombre de classes
- Le choix de la délimitation des classes
- Le choix de la représentation graphique

Comme le titre l'indique, l'objet de ma réflexion ici sera *le troisième choix*. Le choix de la représentation graphique qui est une question de symbolisation. A cet égard, le travail de Jacques Bertin a eu une immense importance.

Pour déterminer la méthode de symbolisation, Robinson (1995) affirment qu'il est nécessaire de distinguer ces deux constituants :

- le choix du niveau de mesure
- le choix des variables visuelles

et de comprendre :

- le rapport entre ces deux constituants

L'objet de cette présentation est d'analyser un problème fondamental associé à ce dernier point : le rapport entre le choix du niveau de mesure et le choix des variables visuelles.

Le terme *niveau de mesure* montre la manière dont les données sont organisées. Autrefois, la manière ordinaire d'organiser les données était soit qualitative, soit quantitative. Cependant, cette dichotomie n'était pas suffisamment détaillée et donc peu satisfaisante puisqu'elle créait des obstacles au progrès scientifique. Afin d'éviter ces obstacles, Stevens (1951) a subdivisé le système de mesure[1] en quatre : *nominal*, *ordinal*, *interval* et *ratio*. Depuis, ce système a été adopté dans plusieurs disciplines telles que la géographie et la cartographie, surtout dans les textes anglo-saxons.

Selon les textes écrits par Bertin, il existe une autre façon d'organiser les données. Bertin emploie un système d'organisation divisé en trois. Par conséquent, des problèmes peuvent en découler car ces deux systèmes d'organisation ne sont pas équivalents.

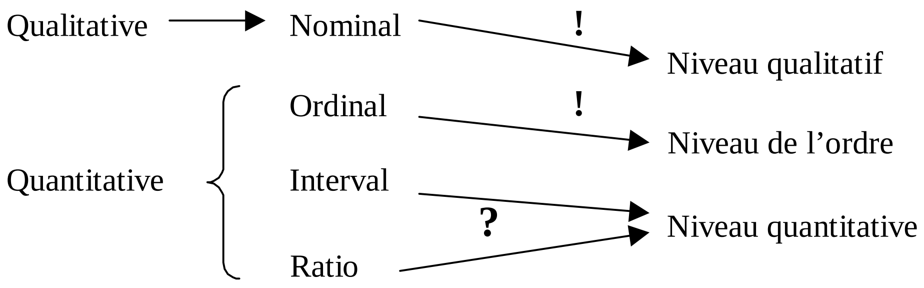


Figure 1. Différents systèmes de mesure et incompatibilité entre les façons divisées en quatre et en trois.

---

[1] Le concept du système de mesure est quelque peu trompeur car on peut difficilement affirmer que le niveau nominal serait mesuré.

Cette incompatibilité devient un problème quand les textes écrits utilisant une division en quatre cherchent à créer des règles de symbolisation fondées sur la Sémiologie Graphique. Celle-ci est adaptée à une division en trois. Par conséquent, les règles de la représentation graphique sont obscures. Cette présentation a pour objectif de montrer ces problèmes et propose une façon de les résoudre. Elle est constituée de deux parties : l'ajout du *niveau manquant* et le *déballage* du terme quantitatif.

Le niveau manquant

Selon MacEachren (1995)

> … he [Bertin] did not distinguish between interval and ratio levels (p: 270).

Apres avoir mené une analyse minutieuse de la manière dont Bertin utilise les termes *niveau de l'ordre* et *niveau quantitatif,* il semble évident qu'ils s'accordent respectivement avec des niveaux *ordinal* et *ratio*. Ainsi, je suis en désaccord avec MacEachren et je dirai que le niveau *interval* <u>manque</u> dans la conceptualisation de Bertin.

Est-ce que le niveau *interval* est nécessaire ?

Selon Robinson et autres, la distinction entre *interval* et *ratio* n'a pas de sens dans la symbolisation cartographique.

> In both instances a range is being displayed, and from the point of view of representation, it is immaterial whether or not the scale begins at an arbitrary zero (Robinson 1984: 110).

Donc, la distinction entre *interval* et *ratio* est importante seulement pour celui qui utilise et interprète les cartes et est sans importance pour les cartographes qui font la symbolisation. Je suis en désaccord avec cela.

Pour répondre à la question si le niveau *interval* est nécessaire, il faut que nous entamions l'étude concernant le contraste de ce niveau avec le niveau *ratio*. Au niveau du *ratio* on est capable de définir des proportionnalités, c'est-à-dire de reconnaître qu'un élément graphique représente une magnitude, par exemple double ou triple, par rapport à un autre élément graphique. Une telle induction est impossible au niveau *interval*, car la variable possède un zéro défini arbitrairement. Les formulations de proportionnalité sont possibles seulement si les variables ont un niveau zéro absolu.

Une autre manière pour distinguer *interval* et *ratio* se trouve chez Kraak et Ormeling (1996) qui ne distinguent pas les points zéro, mais leurs qualités pour mesurer *differences*. Ils utilisent un plan pour la symbolisation divisé en quatre, ce qu'ils appellent *perceptual characteristics* et qui correspond aux niveaux de mesure suivants :

∉# difference in quality - nominal scale
∉# difference in order - ordinal scale
∉# difference in distance - interval scale
∉# difference in size (proportions) - ratio scale

Utilisant cette définition pour le terme *interval,* je préfère le remplacer par le terme *range-graded*

> range-grading acts much like interval scale measurements (Robinson et al 1995:273)

ce qui correspond au niveau *interval*. Pour illustrer le concept d'un niveau *range-graded*, Robinson et autres ont pris l'exemple de quatre groupes de revenus de familles qui sont les suivants :

“mois de $10,000”
“$10,000 à $30,000”
“$30,000 à $50,000”
“plus de $50,000.”

Quelle sorte de *differences* pouvons-nous extraire de ce niveau, *range-graded* ? Nous sommes en mesure d’affirmer qu’il y a une différence de revenu de $40,000 ou plus entre les familles économiquement les plus faibles et les familles appartenant à la catégorie aux revenus les plus élevés. On ne peut pas atteindre une telle différence pareille dans le niveau *ordinal*, puisque ce niveau n’indique pas spécifiquement une aptitude numérique de la différence.

Déballer le niveau quantitatif

Ajouter le niveau manquant entre le niveau de l’ordre et le niveau quantitatif résout le problème de l’éloignement entre ces deux niveaux. Cependant, nous ne comprenons pas encore la nature dont est composé le niveau quantitatif.

Il me semble que cette nature composée sème le désordre dans la symbolisation graphique. Elle rend obscure l’emploi de la variable visuelle *valeur*. On peut illustrer cela par la manière dont Bertin, et ceci contrairement à Weibel et Buttenfield, utilise la *valeur* dans son schéma pour la symbolisation.

	Forme	Orientation	Couleur	Grain	Valeur	Taille
Associatif	∑	∑	∑	∑		
Dissociatif					∏	∏
Selective		#	#	#	#	#
Order				o	O	O
Quantitatif						Q

Table 1. Schéma de symbolisation selon Bertin (1981: 231).

	Shape	Orientation	Colour	Texture	Value	Size
Associative	∑	∑	∑	∑		
Dissociative					∏	∏
Selective		#[2]	#	#	#	#
Ordered				O	O	O
Quantitative					Q	Q

Table 2. Schéma de symbolisation selon Weibel et Buttenfield (1992, figure 3, page 229).

Cette différence est frappante ! Tant Bertin a insisté que seulement la *taille* puisse être utilisée afin de symboliser une variable quantitative, et que

> Le blanc ne pouvant être une unité de mesure pour le gris ou le noir, on ne peut traduire des rapports quantitatifs par la variation de valeur, on ne peut traduire qu’un ordre (Bertin 1967 :48).

Cependant, Weibel et Buttenfield toléreront aussi que la *valeur* puisse symboliser des quantités. Pourquoi ? Selon moi, Weibel et Buttenfield peuvent présenter un tel schéma de la symbolisation et encore être, selon Bertin, en raison de la nature composée du niveau quantitatif. Le terme, donc, a besoin d’être *déballé*.

Après avoir lu soigneusement Bertin, on découvre qu’il subdivise le niveau quantitatif en deux parties : *ratio* et *quantifie absolu*.

---

[2] Weibel et Buttenfield référer Bertin (1983b) qui est une tradiction de: 'Dans les trois impplantations la forme n'est pas sélective, ni l'orientation en implantation zonale' (Bertin 1967:67).

> Graphics separates quantities into two types … *totals per object* and *ratios* between totals per object. The totals are also called “absolute quantities” (Bertin 1981:190).

Un exemple typique du terme *ratio* est la *densité*, par exemple la densité de population. Pour éviter la confusion avec le niveau *ratio* utilisé par les anglo-saxons, le terme *densité* remplacera le terme ratio. De plus, ce terme est aussi employé par Bertin comme un synonyme de ratio.

Donc, ayant déballé le quantitatif en *densité* et *quantité absolue*, nous découvrons que Bertin permet que seule la *taille* représente la *quantité absolue* tandis que la valeur peut représenter des densités. Voilà, un exemple :

	A	B	
P	3	12	Population
S	1	6	Superficie
P/S	3	2	Densité de population

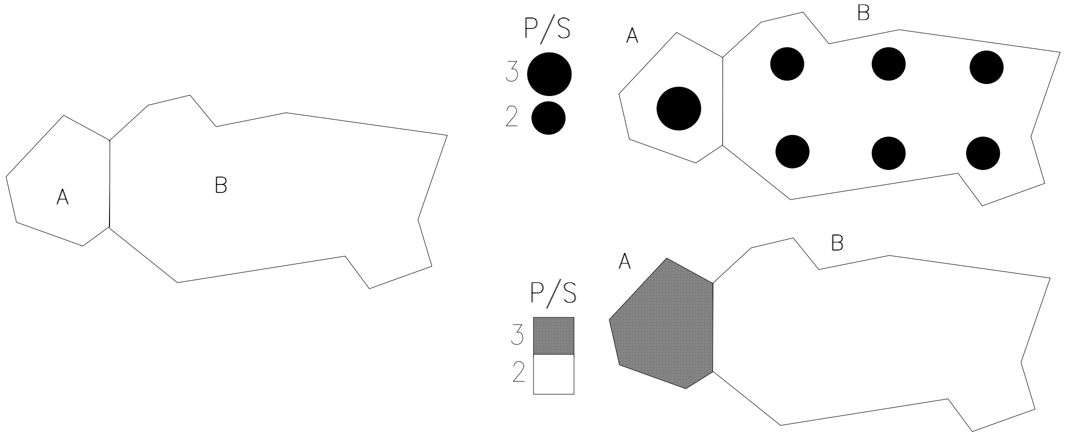


Figure 2. La représentation de densité en utilisant une grille de point symboles (taille) et valeur (d’âpres Bertin 1981:191).

Cela peut expliquer les différents emplois de la *valeur* dans les schémas de la symbolisation qu’utilise Bertin contrairement à Weibel et Buttenfield.

Ainsi on aboutit à ces cinq niveaux :

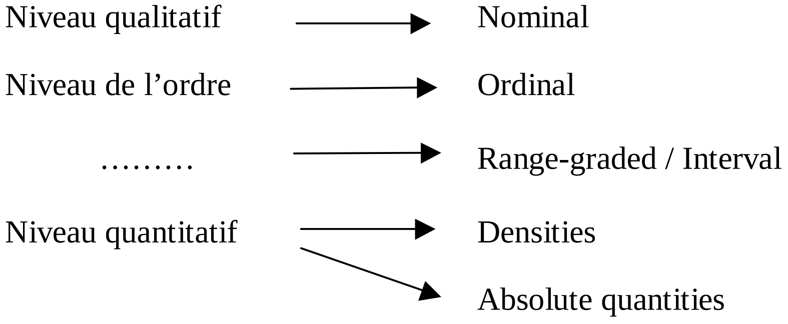


Figure 3. La cause de ces cinq niveaux est le fait d’ajouter le niveau *range-graded* et le déballant de niveau de quantitatif.

Le but dans lequel on fait une carte

Ces niveaux, ils sont tous nécessaires ? Cela dépend avant tout dans quel but la carte est faite. Afin d’arriver à ce but, il faut qu’on choisisse un niveau de mesure convenable. Afin de distinguer deux objectifs principaux que poursuit l’élaboration des cartes, j’utilise les concepts *graphic information processing* et *graphic communication* (Bertin 1981). On voit ces deux objectifs aussi par le biais d’autres concepts.

Bertin (1981)	Muehrcke (1990)	MacEachren (1995)
Graphic information processing	Geographic thinking	Visualisation tool
Graphic communication	Geographic illustration	Communication device

Table 3. Les deux intentions de cartes exprime avec les concepts diffèrent.

Ainsi, en utilisant ces deux objectifs, les variables visuelles et les cinq niveaux de mesure, on arrive au schéma suivant qui constitue une proposition pour une base des règles graphiques:

			Form	Orientation	Colour	Texture	Value	Size
commu-nication	visualisation	Associative	■	■	■	■		
		Dissociative					■	■
		Selective		■	■	■	■	■
		Ordinal				■	■	■
		Interval/range-graded					■	■
		Densities/derived ratio					■	■
		Absolute values						■

Table 4. Les variables visuelles, le niveau de mesure et l’objectif lors de la réalisation des cartes.

Conclusion

Nous ne serions jamais à même d’établir un schéma définitif pour la Sémiologie Graphique. Simplement, parce que, tout comme les autres langues, la langue graphique change et évolue. Pour la Sémiologie Graphique on peut voir cela à travers les résultats obtenus par des cartographes nord américains qui, depuis les 30 dernières années, ont avancé le concept de variables visuelles additionnelles (Morrison 1974, MacEachren 1995). En d’autres termes, ils ont élaboré les données de l’axe horizontal. Dans cette présentation, j’ai élaboré les données de l’axe vertical, c’est-à-dire l’organisation des données qui se situent dans la première colonne du schéma. Tout au long des deux axes, il y a plusieurs problèmes cruciaux pour la réalisation de règles de base de symbolisation.

Dans l’article et dans cette présentation, j’ai eu l’intention de montrer l’incompatibilité entre les niveaux d’organisations proposés par Bertin, et ceux utilisés par les statisticiens et les cartographes anglo-saxons. Dans une optique d’élaboration de règles de base pour un système d’expert destiné à la réalisation de cartes statistiques, il est important que les points qui sont obscures à cause de cette incompatibilité, soient supprimés. Cette suppression est accomplie en ajoutant le niveau *range-graded* et en déballant le terme quantitatif de *densité* et de *quantifie absolu*. J’espère que cela participera à la formalisation de règles de base pour faire la symbolisation graphique.

Références

Baudouin, Axel 1987 : *Klasseinndeling i kartografi* Kart og Plan, No 3, pp. 321-326
Bertin, Jacques 1967: *La Sémiologie graphique*. Paris, Mouton.
Bertin, Jacques 1981: *Graphics and Graphic Information Processing*. Walter de Gruyter.
Bertin, Jacques 1983: *A New Look at Cartography*. In *Graphic Communication and Design in Contemporary Cartography*. Progress in Contemporary Cartografy, Volume 2, Ed. D.R. Fraser Taylor, John Wiley & Sons.
Bertin, Jacques 1983b: *Semiology of Graphics*. Gauthiers-Villars, Paris.
Kraak, M.J. and Ormeling F.J. 1996: *Cartography – Visualisation of spatial data*. Longmann.

MacEachren, Alan M. 1995: *How Maps Work – Representation, Visualization, and Design*. The Guilford Press

Morrison, Joel L. 1974: *A theoretical framework for cartographic generalisation with the emphasis on the process of symbolisation*. International Yearbook of Cartography, 14, 115-127.

Muehrcke, Phillip C. 1990: *Cartography and geographical information system.* Cartography and geographical information system, 17(1), 7-15.

Robinson Arthur H., Sale Randall D., Morrison Joel L., Muehrcke Phillip C. 1984: *Elements of Cartography*, John Wiley & Sons. 5th edition.

Robinson Arthur H., Morrison Joel L., Muehrcke Phillip C., Kimerling A. Jon, Guptill Stephen C. 1995: *Elements of Cartography*, John Wiley & Sons. 6th edition.

Stevens, S. S. 1951: *Handbook of Experimental Psychology*. John Wiley & Sons.

Weibel, Robert and Buttenfield, Barbara P. 1992: *Improvement of GIS graphics for analysis and decision-making*. Int. J. Geographical Information System, vol. 6, no. 3, 223-245.

The third choice …

Jan Ketil Rød

Abstract

30 years ago, Jacques Bertin published *Sémiologie Graphique* (Bertin, 1967). This theory has become fundamental for cartographic education and for applied cartography. I will not review in this paper the prolific employment of the graphic semiology, but rather draw attention towards an inconsistency between the organisation level used by Bertin and the measurement levels used by Anglo-American statisticians and cartographers. As the former has three levels and the latter four, it confuses an approach in elaborating a rule base for an expert system accommodated for statistical mapping. A framework for understanding the issue is provided by reviewing the obvious inconsistencies between the two systems and a possible solution is proposed. This reflection results in a symbolisation schemata applicable as a rule base for the two functions of the map: those used as a visualisation tool and those used as a communication device.

Introduction

This paper is a synthesis of the article written for the occasion of the thirtieth anniversary of Graphic Semiology and deals with symbolisation rules applied for statistical mapping. The decisions necessary for such maps to be realised, as recognised by Baudouin, can be grouped into the following three issues:

- The choice of the number of classes,
- The choice of the method applied for determining class limits, and
- The choice of graphic rendering

As the title of this paper suggests, it is the third choice which will be considered here. The choice of graphic rendering is a question of symbolisation. In this respect, the work of Jacques Bertin has had an immense importance.

To be able to perform symbolisation, Robinson et al (1995) claims that it is necessary to recognise its two components:

- the choice of the level of measurement

⊄# the choice of the visual variables

and to understand the relationship between these two components.

The objective of this presentation is to elaborate an essential problem associated with the last point: the relationship between the choice of the level of measurement and the choice of the visual variable.

The term *measurement level* denotes how data is organised. It used to be common to organise data into either qualitative or quantitative categories. This dichotomy, however, was not recognised as not sufficiently detailed and thus unsatisfactory since it placed restrictions on scientific advancement. To avoid these restrictions, Stevens (1951) subdivided the dichotomy further into four levels: *nominal, ordinal, interval* and *ratio*, which have since been adopted by several disciplines including geography and cartography, especially in Anglo-Saxon literature.

According to the writings of Bertin, there is another way of organising the data. Bertin applies a three-levelled organising system. Consequently, since these two organising systems cannot simply be equated difficulties are created.

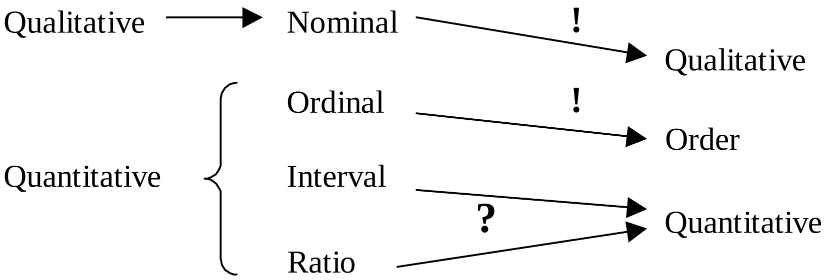


Figure 1. Different measurement systems – the two systems, which are subdivided into four and three levels respectively, are incompatible.

This incompatibility becomes a problem when the texts operating with a four-levelled measurement system attempt to make symbolisation rules based on Bertin’s theory of graphic semiology, which is adapted for a three-levelled organisation system. Accordingly, the rules for graphic representation become obscured. This paper tries to illustrate the problem and proposes an approach to solving this. The approach consists of two parts: adding *the missing level* and *unpacking* the term quantitative.

The missing level

According to MacEachren (1995)

> … he [Bertin] did not distinguish between interval and ratio levels (p: 270).

After a closer examination of how Bertin used the term order and quantitative, it seems obvious that they match ordinal and ratio levels respectively. Thus, I would put it differently than MacEachren by stating that the interval level is missing in Bertin's conceptualisation.

Is the interval level necessary?

According to Robinson, the distinction between interval and ratio is irrelevant for the cartographic symbolisation.

> In both instances a range is being displayed, and from the point of view of representation, it is immaterial whether or not the scale begins at an arbitrary zero (Robinson et al, 1984: 110).

Thus, the distinction between interval and ratio is only important for the map user's interpretation and not for the mapmaker's symbolisation. I would disagree.

In order to answer whether or not the level *interval* is necessary, it is necessary to compare how it is defined compared to the ratio level. At the ratio level one is able to define proportions, which means to be able to ascertain that one graphic element represents a magnitude, for instance, double or triple than another graphic element. Such an induction is impossible at the interval level since a variable at this level has an arbitrary defined zero. Statements of proportions are only possible if the variables' units have an absolute zero.

Another way of differentiating between interval and ratio is found in Kraak and Ormeling (1996), who do not differentiate in terms of their point of zero, but their ability to measure differences. They use a four-levelled symbolisation schema in which four *perceptual characteristics* correspond to the levels of measurements. These are as follows:

- difference in quality - nominal scale
- difference in order - ordinal scale
- difference in distance - interval scale
- difference in size (proportions) - ratio scale

Using this definition, however, I would prefer using the term *range-graded* instead of interval, since:

> range-grading acts much like interval scale measurements (Robinson et al 1995:273)

Consequently, range-graded corresponds with the interval level. To illustrate the concept *range-graded* level, Robinson et al provide the following example with the four groups of family incomes:

“less than $10,000”
“$10,000 to $30,000”
“$30,000 to $50,000”
“above $50,000.”

What kind of *differences* can we extract from the level *range-graded*? It is possible to say that a family belonging to the lower income bracket has a difference in income of $40,000 or more in comparison with a family in the upper income bracket, hence a difference in distance. Such a difference in distance is not obtainable at the ordinal level, since this level does not indicate any specific numerical magnitude of difference.

Unpacking the term quantitative

Adding the missing level between ordinal and quantitative solves the problem of the considerable leap between these two levels. However, the compounded nature of the term quantitative is not yet clear.

This compound nature seems to be the source of confusion for the graphic symbolisation. It obscures the use of the visual variable *value*, which can be illustrated by the way Bertin versus Weibel and Buttenfield apply *value* in their symbolising schema.

	Shape	Orientation	Colour	Texture	Value	Size
Associative	Σ	Σ	Σ	Σ		
Dissociative					Γ	Π
Selective		#	#	#	#	#
Ordered				o	O	O
Quantitative						Q

Table 1. Symbolisation schema according to Bertin (1981: 231).

	Shape	Orientation	Colour	Texture	Value	Size
Associative	Σ	Σ	Σ	Σ		
Dissociative					Π	Π
Selective		#[1]	#	#	#	#
Ordered				O	O	O
Quantitative					Q	Q

Table 2. Symbolisation schema according to Weibel and Buttenfield (1992, figure 3, page 229). The shade of grey is added in order to emphasise the difference from Bertin's equivalent symbolisation scheme (see Table 1).

The difference is striking! While Bertin claimed that only *size* could be used in order to symbolise a quantitative variable, and that:

> Since white cannot provide a measuring unit for gray or black, quantitative relationships cannot be translated by a value variation. Value can only translate an order (Bertin 1983b: 48).

However, Weibel and Buttenfield also allow *value* to symbolise quantities. Why? I believe that Weibel and Buttenfield, although presenting such a symbolisation schema (Table 2), may still be in accordance with Bertin because of the compounded nature of the term quantitative. The term needs therefore to be *unpacked*.

After a close reading of Bertin, one discovers that he divides the term quantitative into two parts: ratio and absolute quantities.

> Graphics separates quantities into two types … *totals per object* and *ratios* between totals per object. The totals are also called "absolute quantities" (Bertin 1981:190).

One typical example of the term *ratio* is *density*, for instance, population density. To avoid confusion between the level *ratio* used in Anglo-Saxon literature, the term *density* will in the following replace what Bertin calls ratio. In addition, density is also a term used by Bertin as a synonym for what he calls ratio.

Thus having unpacked quantitative into *density* and *absolute quantities*, it is evident that Bertin allows only *size* to represent *absolute quantities* while value may be applied to represent densities. Here, an example:

---

[1] Weibel and Buttenfield refer to Bertin (1983b) or more precisely: 'For all three implantations – point, line, and area – shape is not selective; nor is orientation when represented by area' (Bertin, 1983b: 67).

	A	B	
P	3	12	Population
S	1	6	Surface area
P/S	3	2	Population density

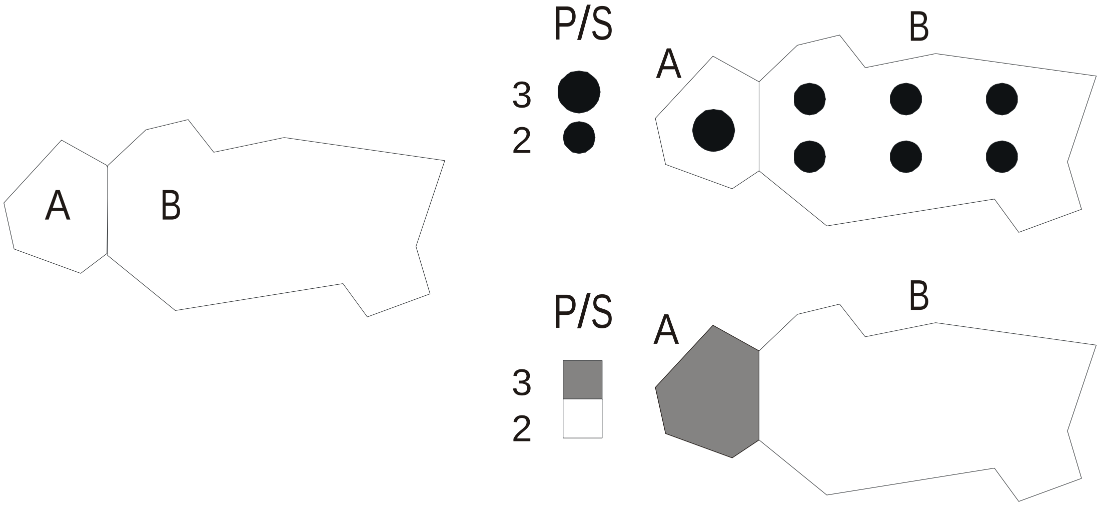


Figure 2. Representation of density by using size variation (in a grid of point symbols) and value variation (from Bertin, 1981: 191).

This possibly explains the differences between the symbolisation schema according to Bertin and Weibel and Buttenfield respectively.

Thus, the end result is five levels:

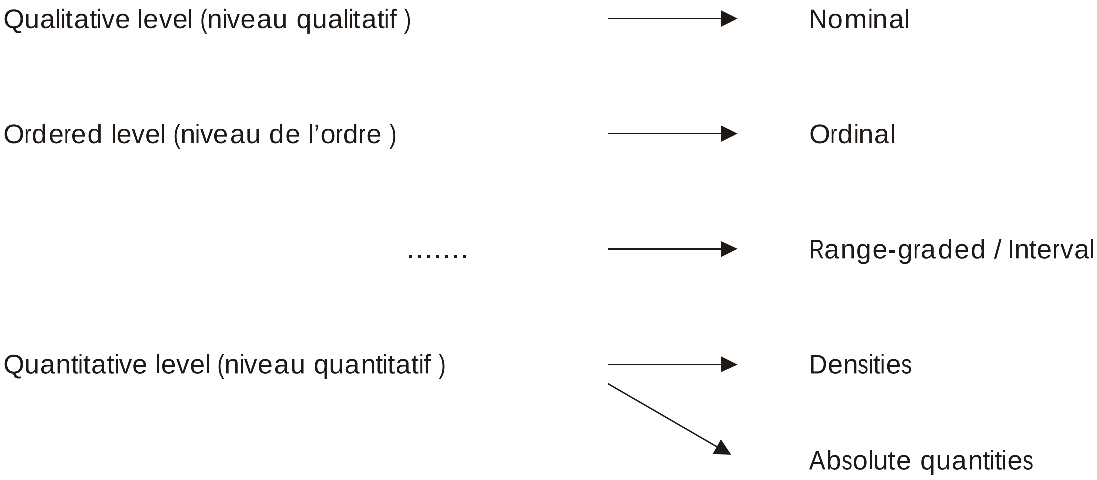


Figure 3. These five levels are the result of adding the level *range-graded* and unpacking the quantitative level.

The reason for why one makes a map

Are all these levels necessary? Whether all levels are necessary depends, firstly, on the intended use of the map. I will use the notions *graphic information processing* and *graphic communication* (Bertin, 1981) to distinguish between two basic intentions of map-use underpinning the making of a map. These two intensions of map-use can be expressed additionally with matching concepts:

Bertin (1981)	Muehrcke (1990)	MacEachren (1995)
Graphic information processing	Geographic thinking	Visualisation tool
Graphic communication	Geographic illustration	Communication device

Table 1. The two intensions if map use expressed by different concepts.

Thus by combining the two uses of maps, five levels of measurement and visual variables, the following schema results, which constitute a proposal for a graphic rule base:

			Shape	Orientation	Colour	Texture	Value	Size
commu-nication	visualisation	Nominative (Associative)	X	X	X	X		
		Nominative (Dissociative)					X	X
		Nominative (Selective)		X	X	X	X	X
		Ordinal				X	X	X
		Interval/range-graded					X	X
		Densities/derived ratio					X	X
		Absolute values						X

Table 6. Two uses of maps, levels of measurement and the visual variables set together to produce a symbolisation scheme.

Conclusion

It seems impossible to arrive at a final scheme for the Semiology of Graphics, simply because, just like any other language, the graphic language alters and develops. The Semiology of Graphics has advanced during the last thirty years, especially according to the results achieved by North-American cartographers with additional visual variables (Morrison, 1974, MacEachren, 1995). To put it another way, efforts by the North-American cartographers have elaborated the concepts along the horizontal axe in Table 4. In this paper, I have considered in detail the concepts on the vertical axe in Table 4. Still, along both axes, it is evident that there are several problems essential for establishing a rule base for symbolisation.

With this paper, I had the intension of indicating the inconsistency between Bertin’s level of organisation and the measurement levels used by the Anglo-Saxon statisticians and cartographers. To provide more clearly defined rules for an expert system applicable for statistical mapping, it is important that the obscurities, produced by this inconsistency, are removed. This is tried by adding the level *range-graded* and by “unpacking” Bertin’s term quantitative into *densities* and *absolute values*. I hope that this will contribute to the formalisation of a rule base applicable for the graphic symbolisation.

Article IV

Rød, J.K. 2001a. How the monosemic graphics go polysemic. *Cartographic Perspectives.* Number 38, Winter 2001: 26-37.

Article IV is not included due to copyright.

Article V

Rød, J.K. 2000. De nauwkeurigheid van geclassificeerde kaarten. *Kartografisch Tijdscrift* 16 (4): 47-51.

The accuracy of classified maps

De nauwkeurigheid van geclassificeerde kaarten

J. K. Rød

thematische kartografie, classificatie van kwantitatieve gegevens, visualisatie | TREFWOORDEN
thematic cartography, classification of quantative data, visualisation | KEYWORDS
cartographie thématique, classification des données quantitatives, visualisation | MOTS-CLÉS

Kwantitatieve ruimtelijke gegevens kunnen worden geclassificeerd om reeds bekende patronen aan anderen over te brengen, of om nog onbekende patronen te ontdekken. Om te voorkomen dat onjuiste ruimtelijke patronen worden overgebracht of ontdekt, zou een foutgevoelige GIS uitkomst kunnen bieden. Dit artikel laat zien hoe een GIS of kartografisch pakket door middel van een eenvoudige visualisatie en een index informatie over de nauwkeurigheid van de geclassificeerde kaart naar de gebruiker kan terugkoppelen. De gebruiker kan daardoor beoordelen in hoeverre een classificatie overeenkomt met de oorspronkelijke gegevens. En hoe beter de classificatie is, hoe geringer de kans is op onjuiste ruimtelijke patronen.

Jan Ketil Rød is AIO bij het Geografisch Instituut, Universiteit Trondheim, Noorwegen (NTNU), [t] + 47 73 59 17 23, [e] jan.ketil.rod@svt.ntnu.no.

Inleiding

Dit artikel behandelt de nauwkeurigheid van een klassenindeling die wordt toegepast op een variabele voor afbeelding in een thematische kaart. Volgens Bertin kan er onderscheid worden gemaakt tussen thematische en topografische nauwkeurigheid, waarbij hij stelt dat thematische nauwkeurigheid belangrijker is. Een onnauwkeurige thematische kaart zal, wanneer deze wordt gebruikt als basis voor besluiten, kunnen leiden tot slechte besluiten, niet door een fout in de geografische positie, maar in de geografische patronen [Bertin, 1983, p. 70]. Om een betere basis voor besluitvorming te creëren, heeft Bertin een serie 'regels' ontworpen om de grafische presentatie te laten corresponderen met de karakteristieken van de data [Bertin, 1981, 1983b]. Zo dienen geordende of kwantitatieve gegevens te worden uitgebeeld door corresponderende visuele variabelen. Tegenwoordig kan men er bijna altijd vanuit gaan dat commerciële GIS- en kartografische pakketten een functie hebben die automatisch een bij het karakter van de gegevens passende visuele variabele toekent aan geclassificeerde data (bijvoorbeeld: klassen met hogere waarden krijgen een donkerder signatuur of grotere symbolen).

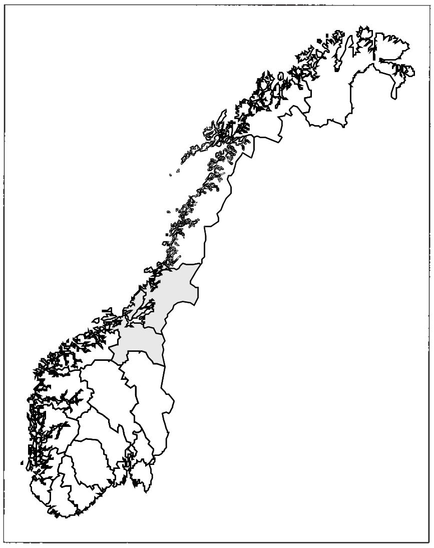
Volgens Baudouin [1987] kan de indeling in klassen voor thematische kaarten in drie componenten worden gesplitst: de keuze van het aantal klassen, de keuze voor de klassengrenzen en de keuze voor een wijze van grafische weergave. Mede dankzij de erfenis van Bertin is de keuze voor een bepaalde weergave al vrij ver in de grafische pakketten doorgevoerd. De eerste twee keuzes daarentegen behoeven nog enige aandacht, met name met betrekking tot de nauwkeurigheid van de classificatie. Deze nauwkeurigheid zal het belangrijkste aandachtspunt van dit artikel vormen.

De relatie tussen nauwkeurigheid en de twee bovengenoemde componenten van de klassenindeling voor thematische kaarten komt in de kartografische literatuur veelvuldig terug [Jenks, 1963, 1977; Jenks & Caspall, 1971; Evans, 1977]. In commerciële pakketten is echter slechts een klein gedeelte hiervan geïmplementeerd. Eén van de methoden om de nauwkeurigheid van een klassenindeling aan te geven, is een index waarin af te lezen is in welke mate de geclassificeerde verdeling overeenkomt met de oorspronkelijke verdeling. De meest voorkomende maat voor deze overeenkomst is de *'Goodness of Variance Fit'* (GVF). Dit artikel behandelt de berekening van de GVF in het kort en zal vervolgens ingaan op de betekenis van deze maat met een voorbeeld van de twee provincies Sør- en Nord-Trøndelag in Noorwegen (figuur 1).

De 'Goodness of Variance Fit'

De GVF is een maat voor de statistische overeenkomst tussen de klassenindeling en de oorspronkelijke data. Hoe groter de homogeniteit binnen de klassen (kleine spreiding) en tevens hoe groter de heterogeniteit tussen de klassen (grote spreiding), hoe beter de classificatie is. De GVF wordt op de volgende manier berekend [Dent, 1996, pp 136]:

Bereken het rekenkundig gemiddelde voor de variabele en tel de gekwadrateerde afwijkingen tussen de observatiewaarden en het gemiddelde bij elkaar op. Het resultaat wordt de SDAM



Figuur 1 - De provincies Sør- en Nord-Trøndelag in Noorwegen.

(Squared Deviation of the Array Mean) genoemd (zie kader, formule 1).

Bereken het rekenkundig gemiddelde voor elke klasse. Tel vervolgens per klasse de gekwadrateerde afwijkingen tussen de observatiewaarden binnen de actuele klasse en zijn rekenkundig gemiddelde op. Voor één klasse wordt dit de SDCMC genoemd. Tenslotte worden alle SDCMC 's bij elkaar opgeteld. Het resultaat wordt de SDCM (Squared Deviation of Class Means) genoemd (zie kader, formule 2).

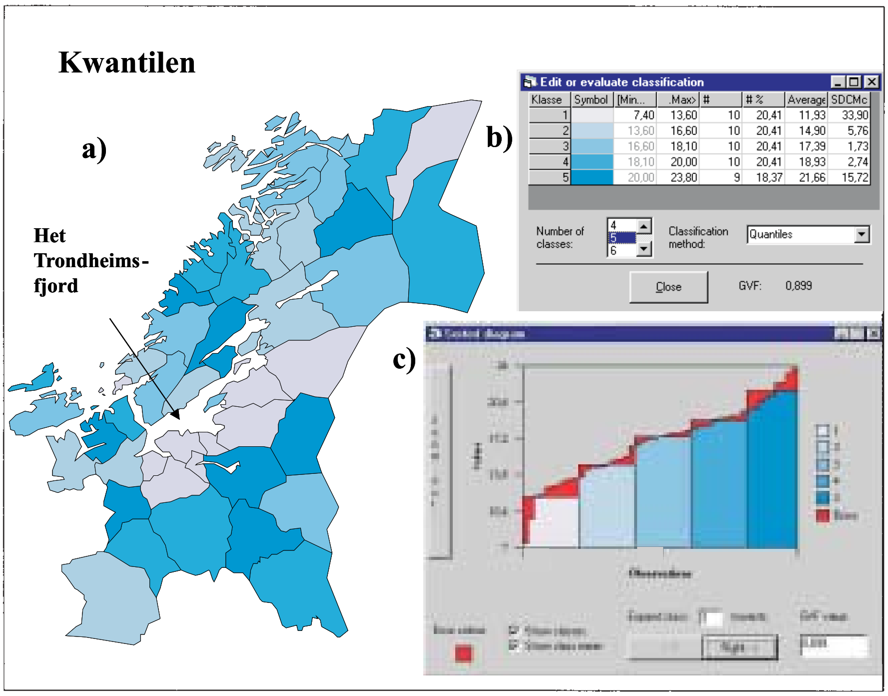
Bereken de GVF als het verschil tussen SDAM en SDCM gedeeld door SDAM (zie kader, formule 3).

$$SDAM = \Sigma (x_i - X)^2$$

$$\begin{array}{l} SDCM \\ \cancel{SDAM} = \Sigma\Sigma (x_{ic} - Z_c)^2 \end{array}$$

$$GVF = \left(\frac{SDAM - SDCM}{SDAM}\right)$$

Figuur 2 - Aandeel van de leeftijdsgroep boven 67 jaar op de totale bevolking per gemeente in Sør- en Nord-Trøndelag (1999). Indeling in vijf klassen volgens de kwantielenmethode.



Klasse	Symbol	[Min...	.Max>	#	# %	Average	SDCMc
1		7,40	13,60	10	20,41	11,93	33,90
2		13,60	16,60	10	20,41	14,90	5,76
3		16,60	18,10	10	20,41	17,39	1,73
4		18,10	20,00	10	20,41	18,93	2,74
5		20,00	23,80	9	18,37	21,66	15,72

Number of classes: 5

Classification method: Quantiles

GVF: 0,899

De theoretische waarden voor de GVF liggen altijd tussen 0 en 1. In het algemeen neemt de GVF toe met het aantal klassen. Voor choropletenkaarten waarin het aantal klassen correspondeert met het aantal eenheden met een unieke waarde in de dataset (klassenloze choropletenkaart) zal de GVF gelijk zijn aan 1 indien de 'indeling' identiek is aan de oorspronkelijke verdeling.

Het voorbeeld

Figuur 2a laat het aandeel zien van de leeftijdsgroep boven 67 jaar op de totale bevolking per gemeente in Sør- en Nord-Trøndelag in 1999. De variabele is ingedeeld in vijf klassen volgens de kwantielenmethode, waarbij het aantal waarnemingen evenredig of bijna evenredig wordt verdeeld over het aantal klassen. In bepaalde commerciële kartografische pakketten wordt deze methode als de prototype-classificatie aangeboden. Een oude observatie lijkt daarom nog steeds te gelden:

> [Kwantielen] *"...have been selected by cartographers wishing to play safe and make sure that some spatial differentiation was portrayed."*
> *[Evans, 1977, p. 107].*

Volgens Evans worden ruimtelijke verschillen vaak versterkt wanneer de kwantielenmethode wordt gebruikt. Dat is ook het geval in figuur 2a. Een geografische differentiëring die niet duidelijk aanwezig is in de data wordt versterkt, terwijl tegelijkertijd wel aanwezige verschillen worden verborgen.

Wanneer een kartografische classificatie verschillen toont die niet of slechts in beperkte mate aanwezig zijn, dan 'liegt' de kaart [Monmonier, 1991]. Liegen met kaarten doet men bewust of onbewust. Om te vermijden dat een gebruiker van kartografische pakketten onbewust liegt, zou de betreffende persoon een terugkoppeling moeten kunnen krijgen waaruit duidelijk wordt hoe goed een gekozen classificatie is. Hierbij kan de GVF behulpzaam zijn. Daarmee kan tevens gevisualiseerd worden in welke mate iedere klasse past bij de verdeling. Figuur 2b laat de GVF en de bijdragen van de verschillende klassen daaraan (uitgedrukt in SDCMC's) zien. De GVF bedraagt 0,899. De klassen 1 en 5 (met SDCMC's van achtereenvolgens 33,9 en 15,72) dragen het meeste bij aan de slechte GVF. Daarmee worden de interne verschillen tussen de gemeenten rondom Trondheim verborgen, terwijl in plaats daarvan een dramatisch contrast tussen enerzijds de gemeenten in het binnenland (klasse

5) en anderzijds de gemeente langs de Trondheimsfjord (klasse 1) wordt getoond. Verder wordt met deze classificatiemethode gesuggereerd dat een aantal gemeenten aan de buitenkant van de provincie een hoog aandeel oude bevolking heeft doordat er tien gemeenten (in plaats van twee zoals in figuur 3) in de laagste klasse vallen.

Het gebruik van de GVF

De theoretisch maximum waarde van de GVF is 1,0 en derhalve zullen classificaties beter zijn naarmate ze een waarde hebben die dichter bij 1,0 ligt. De GVF kan worden gebruikt om een classificatie te optimaliseren [Jenks & Caspall, 1971], en er bestaan algoritmen die er gebruik van maken [Jenks, 1977; Lindberg, 1990]. De zoektocht naar de optimale kaart illustreert datgene waar kartografen zich mee bezig hebben gehouden na ongeveer 1970. Deze periode wordt wel die van het communicatie-paradigma in de kartografie genoemd [MacEachren, 1995]. Binnen dit paradigma werd de kaart gezien als een medium voor communicatie over een bepaalde, en bekende, boodschap. Vanaf ongeveer 1990 is de interesse van de kartografen verschoven naar het visualiseringperspectief. Volgens dit perspectief houdt men zich niet meer bezig met het vinden van de optimale kaart:

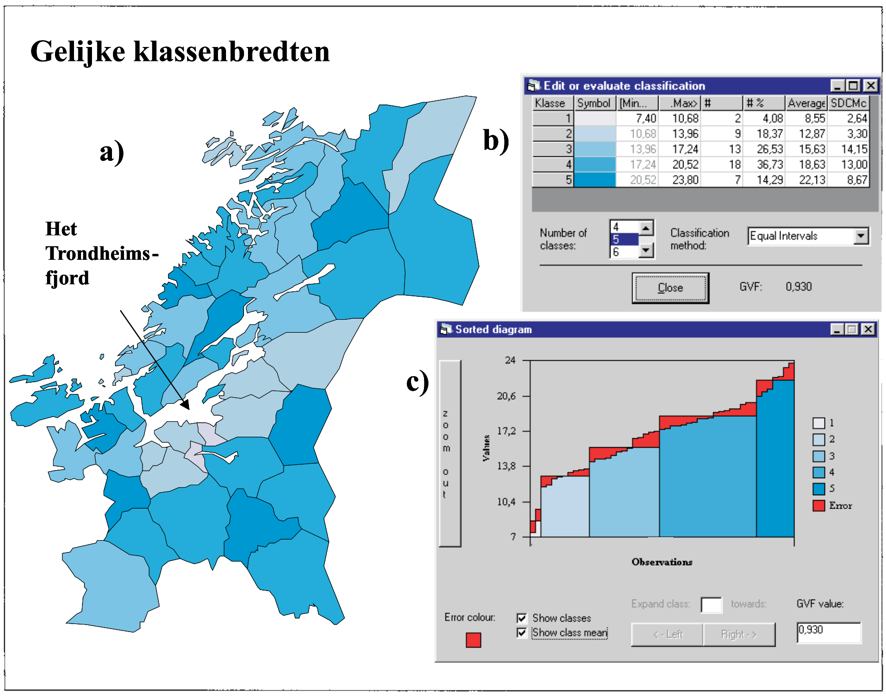
> *"For cartographic visualization the message is unknown and, therefore, there is no optimal map! The goal is to assist an analyst in discovering patterns and relationships in the data."*
> *[MacEachren & Ganter, 1990, p. 65].*

Naar mijn mening is de zoektocht naar het optimaliseren van een geclassificeerde kaart op basis van de GVF echter nog steeds actueel. Met een betere GVF zal de kans afnemen dat verschillen of gelijkenissen die in de dataset aanwezig zijn in de kaart verborgen blijven, of dat juist de verschillen of gelijkenissen die niet in de dataset aanwezig zijn in de kaart naar voren komen. De GVF is derhalve relevant, ongeacht of men deze beschouwt vanuit een communicatie-paradigma of vanuit een visualiseringperspectief. Wanneer de kaart wordt gebruikt als een medium voor *communicatie* geldt: hoe beter de GVF, hoe kleiner de kans is dat de kaart een leugen over zal brengen. Wanneer de kaart wordt gebruikt als een gereedschap voor *visualisering* geldt: hoe beter de GVF, hoe groter de kans dat een patroon dat een analist eventueel in de kaart ontdekt daadwerkelijk bestaat. Indien de kaart een basis voor besluitvorming vormt, wordt het eens te meer duidelijk hoe belangrijk de juiste representatie is. De berekening van de GVF om een indeling in klassen te optimaliseren, of het gebruik van de GVF als hulpmiddel bij het kiezen tussen verschillende klassenindelingen [Declercq, 1995] is daarom nog steeds relevant.

Een op de tweehonderdduizend

De kaart in figuur 2a is slechts één van de 194.580 mogelijkheden om deze variabele met 49 eenheden in te delen in 5 klassen. Het is echter lastig om te bepalen of de klassenindeling die gebruikt is voor de kaart in figuur 2a goed is of slecht zonder de maximale bereikbare GVF voor deze dataset te kennen. Daarom zijn met behulp van een *Visual Basic* algoritme voor alle 194.580 mogelijke klassenindelingen de GVF-waarden berekend. Daaruit blijkt dat er een maximale waarde van 0,937 bereikt kan worden. Declercq [1995, pp. 922] heeft een onderzoek gedaan naar de invloed van het aantal klassen op de GVF bij verschillende classificatiemethoden, waaruit naar voren kwam dat de GVF nivelleert bij waarden boven de 0,95 - ook al neemt het aantal klassen nog toe. Hieruit concludeerde hij dat een GVF van 0,95 kan worden beschouwd als de norm voor een goede classificatie. Zo bleek dat voor de classificatie van een variabele met 308 eenheden zes klassen nodig waren om deze GVF te bereiken. Als we deze conclusie volgen is er voor onze dataset geen normatief goede indeling mogelijk wanneer vijf klassen worden gebruikt. Indien het aantal klassen wordt verhoogd tot zes krijgen we een GVF van 0,954 bij indeling van de waarnemingen in gelijke klassenbreedte, en slechts 0,907 bij een indeling volgens de kwantielenmethode.

Voor karteringsdoeleinden dient dus een klassenindeling te worden gekozen met een GVF die de optimale waarde zo dicht mogelijk benadert. In de praktijk van de commerciële pakketten betekent dit dat de classificatiemethode wordt gekozen die de hoogste GVF geeft. Voor onze dataset blijkt - bij vijf klassen - een indeling met gelijke klassenbreedten de beste oplossing te zijn (zie figuur 3).



Edit or evaluate classification

Klasse	Symbol	[Min...	.Max>	#	# %	Average	SDCMc
1	⬜	7,40	10,68	2	4,08	8,55	2,64
2	🟦	10,68	13,96	9	18,37	12,87	3,30
3	🟦	13,96	17,24	13	26,53	15,63	14,15
4	🟦	17,24	20,52	18	36,73	18,63	13,00
5	🟦	20,52	23,80	7	14,29	22,13	8,67

Number of classes: 5

Classification method: Equal Intervals

Close

GVF: 0,930

Error colour: 🟥

☑ Show classes

☑ Show class mean

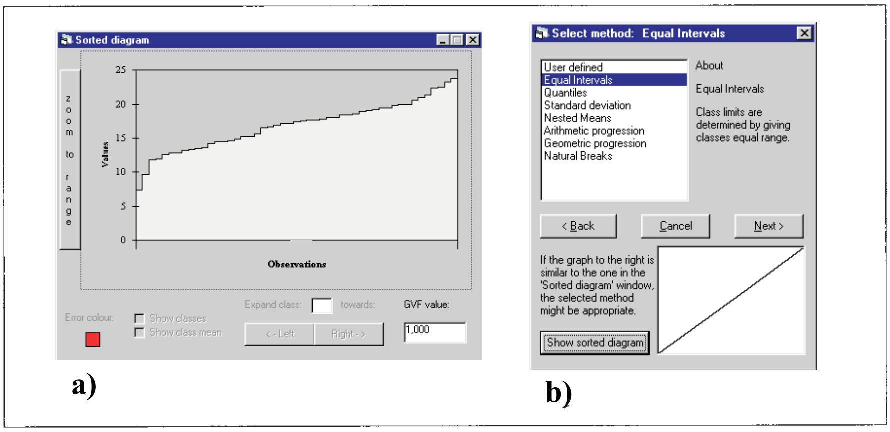
Expand class: towards:

< - Left Right - >

GVF value: 0,930

Figuur 3 - Aandeel van de leeftijdsgroep boven 67 jaar op de totale bevolking per gemeente in Sør- en Nord-Trøndelag (1999). Indeling in vijf klassen met gelijke klassenbreedten.

Een visuele vergelijking van de curve van de oorspronkelijke



Class limits are determined by giving classes equal range.

If the graph to the right is similar to the one in the 'Sorted diagram' window, the selected method might be appropriate.

Figuur 4 - Een visueel hulpmiddel om een geschikte classificatiemethode te kiezen. Links staat de oorspronkelijke verdeling van de dataset, rechts een dialoogvenster dat de ideale curve voor een klassenindeling met gelijke klassenbreedten weergeeft.

verdeling (figuur 4a) en de vorm die de verdeling idealiter zou moeten hebben om geschikt te zijn voor een bepaalde klassenindeling (de theoretische curve, zie figuur 4b) kan de gebruiker van kartografische pakketten helpen bij de keuze van een toepasselijke classificatiemethode. Wanneer de curve van de waarnemingen in een grafiek bijvoorbeeld gelijk is aan een lineaire functie is een indeling met gelijke klassenbreedten de juiste keuze [Ormeling & Kraak, 1987, pp 78, figuur 3-30].

Soms is de curve van de oorspronkelijke verdeling niet eenduidig visueel te interpreteren. In figuur 5 zijn drie curves voor kwantitatieve reeksen weergegeven: (a) lineaire, (b) normale en (c) progressief convexe verdeling van de gegevens. Deze verdelingen zijn achtereenvolgens geschikt voor: (a) gelijke klassenbreedten, (b) standaard deviatie als classificatie-eenheid, (c) meetkundige reeks. Visuele inspectie laat zien dat al deze curves min of meer vergelijkbaar zijn met de oorspronkelijke verdeling (figuur 4a). Bij dergelijke twijfel kan de GVF gebruikt worden om uit te maken welke methode het beste past bij de oorspronkelijke verdeling. De GVF's voor de drie classificatiemethoden bij een indeling in vijf klassen zijn:

Gelijke klassenbreedten:	0,930
Standaard deviatie:	0,924
Meetkundige reeks (progressief convex):	0,923

Alhoewel alle indelingen een hoge GVF geven, gaat de voorkeur uit naar de methode met de beste waarde: gelijke klassenbreedten.

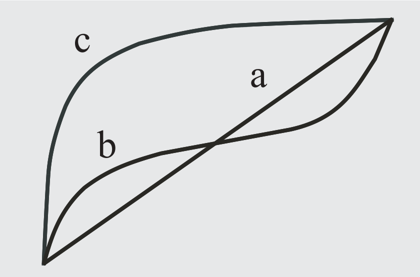
Visualisering van de onnauwkeurigheid van de classificatie

Een losstaand cijfer tussen 0 en 1 zal niet noodzakelijkerwijs begrijpelijk zijn voor de gebruiker van een kartografisch pakket. De GVF dient daarom vergezeld te gaan van een grafische weergave die zijn betekenis duidelijk maakt. Dat kan bijvoorbeeld door middel van een gesorteerd staafdiagram dat de verdeling van de variabele laat zien met de klassengrenzen en de klassengemiddelden, en waarin met een bepaalde kleur wordt getoond hoe de verschillende eenheden van deze gemiddelden afwijken, zoals in de figuren 2c en 3c. In deze figuren is per klasse de afwijking van het klassengemiddelde in rood weergegeven. De oppervlakken van de afwijkingen per klasse vormen bij elkaar de som van de afwijkingen over alle klassen, ofwel de classificatiefout. De relatieve bijdrage van elke klasse aan de classificatiefout is dus eveneens gevisualiseerd. Dit is overigens niet hetzelfde als de SDCMc-waarde (zie figuren 2b en 3b), want dat is de gekwadrateerde som van de afwijkingen binnen een klasse. Waarschijnlijk zal het toch makkelijker zijn om een GVF waarde te begrijpen, als men ziet dat er een samenhang is tussen de grootte van de gerepresenteerde classificatiefout en de GVF. Als het oppervlak van de classificatiefout kleiner wordt - en de classificatiefout dus afneemt - komt de GVF dichter bij 1 te liggen. Waarschijnlijk is het ook makkelijker om te begrijpen welke klasse de grootste bijdrage levert aan de classificatiefout, aangezien de oppervlakte van de afwijking van deze klasse het grootste is.

Het is niet ongewoon om commerciële pakketten te vinden waarbij de klassengrenzen gemakkelijk interactief door de gebruiker zijn aan te passen. In het pakket Descartes [Andrienko & Andrienko, 1998] worden de klassengrenzen aangeven in een 'number line plot' ofwel een spreidingsdiagram. Klassengrenzen zijn er te veranderen met bewegingen van de muis. Als men hetzelfde zou kunnen doen in een gesorteerde staafdiagram zou de gebruiker zelf kunnen ervaren dat verandering van de klassengrenzen de juistheid van de klassenindeling beïnvloedt, omdat de oppervlakte van de classificatiefout verandert.

Conclusie

Voor geclassificeerde kaarten geldt dat de waarschijnlijkheid dat er geografische patronen verborgen blijven in kaarten met een lage GVF groter is dan in kaarten met een hoge GVF. Als men een gebruiker van pakketten waarmee thematische kaarten worden gemaakt een GVF geeft en bovendien de gevisualiseerde bijdrage van elke klasse aan de classificatiefout, zal dit leiden tot een beter inzicht in de nauwkeurigheid van de klassenindeling. Tevens zou de GVF in software kunnen worden gebruikt om die classificatiemethode die de hoogste waarde oplevert als standaard optie aan te bieden. Uiteindelijk zal dit ten goede kunnen komen aan het overbrengen van reeds bekende gegevens aan derden of de exploratie van nog onbekende gegevens.



Figuur 5 - Curves van drie kwantitatieve reeksen: a. lineair, b. normaal, c. meetkundig (progressief convex).

Literatuur

- Andrienko, G.L. & N.V. Andrienko (1998), *Interactive maps for visual data exploration*. Gepresenteerd op de bijeenkomst van de ICA Commission on Visualization, Warsaw, Polen, *http://allanon.gmd.de/and/icavis/* (4 april 2000).
- Baudouin, A. (1987), Klasseinndeling i kartografi. *Kart og plan*, 47(3), pp. 321-326.
- Bertin, J. (1981), *Graphics and graphic information processing*. Berlin: Walter de Gruyter.
- Bertin, J. (1983), A new look at cartography. In: Taylor, D.R.F. (ed.), *Graphic communication and design in contemporary cartography*. Chichester: Wiley, pp. 69-86.
- Bertin, J. (1983b), *Semiology of graphics; diagrams, networks, maps*. Madison: University of Wisconsin Press.
- Declercq, F.A.N. (1995), Choropleth map accuracy and the number of class intervals. In: *Proceedings 17th International Cartographic Conference*, Barcelona: ICA, pp. 918-922.
- Dent, B.D. (1996), *Cartography: Thematic map design*. Dubuque: Brown.
- Evans, I.S. (1977), The selection of class intervals. *Transactions of the Institute of British Geographers*, 2(1), pp. 98-124.
- Jenks, G.F. (1963), Generalization in statistical mapping. *Annals of the Association of American Geographers*, 53, pp. 15-26.
- Jenks, G.F. (1977), Optimal data classification for choropleth maps. *Occasional Paper No. 2*, Department of Geography University of Kansas Lawrence, Kansas.
- Jenks, G.F. & F.C. Caspall (1971), Error on choroplethic maps: definition, measurement, reduction. *Annals of the Association of American Geographers*, 61(2), pp. 217-244.
- Ormeling, F.J. & M.J. Kraak (1993), *Kartografie: Visualisatie van ruimtelijke gegevens*. Delft: Delftse Universitaire Pers.
- Lindberg, M.B. (1990), A Turbo Pascal unit for optimal partitions. *Computers and Geosciences*, 16(5), pp. 717-732.
- MacEachren, A.M. (1995), *How maps work: representation, visualization and design*. New York: Guilford Press.
- MacEachren, A.M. & J.H. Ganter (1990), A pattern identification approach to cartographic visualization. *Cartographica*, 27(2), pp. 64-81.
- Monmonier, M. (1991), *How to lie with maps*. Chicago: University of Chicago Press.

Summary

J.K. Rød - The accuracy of classified maps.

Keywords: thematic mapping, classification of quantitative data, visualisation

Classified choropleth maps are used both in order to communicate particular known geographies and to explore unknown geographies. In both cases, it is important that information about the classification accuracy is given. Using an accurate classified map reduces the probability that the map enhance an impression of geographical differences that only to a small extent is present in the data, or that the map hides differences that indeed is present in the data. If users of a GIS or cartographic software are informed about how accurate the current classification is, than they are better equipped in making maps functioning as sources for geographical knowledge rather than geographical nonsense.

Résumé

J.K. Rød - L'exactitude des cartes classifiées.

Mots-clés: cartographie thématique, classification des données quantitatives, visualisation

Les cartes choroplèthes servent tout autant à décrire qu' à découvrir les structures de l'espace. Dans les deux cas il est important de fournir des indications sur la marge d'erreur des classifications utilisées. L'utilisation d'une classification statistiquement précise réduit le risque de créer des différences géographiques qui ne seraient pas présentes dans les données où d'effacer des différences réelles qui seraient présentes. En informant les utilisateurs des SIG ou de programmes de cartographie sur la précision des classifications qu'ils utilisent on contribuera à leur fournir des outils capables de créer des connaissances plutôt que des non-sens géographiques.

The accuracy of classified maps

Jan Ketil Rød

Quantitative geographical data may be classified to present already known patterns or to discover new patterns. An error sensitive GIS can offer a solution in order to avoid that the geographic pattern, which is presented or discovered, is false. This article demonstrates how, with the use of a plain visualisation and an index, such an error sensitive GIS or mapping package might provide the user with feedback about the classified map’s accuracy. By these means, the user may judge the extent to which a classification corresponds with the original data. Accordingly, the more accurate a classification is, the probability of dealing with a false geographical pattern is far less.

Introduction

This article is about the classification accuracy for a variable, represented in a thematic map. According to Bertin it is: ‘The most widespread and serious error, because it leads to wrong decisions, consists of mistaking not the geographical position but the characteristics. To represent the inherent *order* of quantities by a visual *non-order* or *disorder* of the signs is obviously a mistake and therefore gives a false image – in other words, false information’ [Bertin, 1983, p. 70]. To form a better basis for decision making, Bertin developed a set of ‘rules’ to insure that the graphical presentation better corresponds with the characteristics of the data (1981, 1983b). A quantitative variable should be represented, for example, with a corresponding visual variable. In contemporary commercial GIS and mapping packages, one can almost take it for granted that the groupings from a classified quantitative variable is automatically represented with a corresponding visual variable (for instance: groups having higher values are given a darker shade or larger symbol).

According to Baudouin [1987], the process of classification for thematic maps might be broken down into its three components: (1) the selection of number of classes, (2) the selection of class intervals, and (3) the selection of graphic rendering. Thanks to the sustained influence of Bertin’s groundbreaking work, the selection of graphic rendering is well integrated in graphical software packages. However, the two first points need further clarification, particularly with regard to classification accuracy. This issue forms the central concern of this article.

The relation between accuracy and the two above-mentioned components of data classification for thematic mapping is a well-discussed issue within cartographic literature [Jenks, 1963; Jenks & Caspall, 1971; Evans, 1977]. In commercial packages, however, little more than a small part of this knowledge has been implemented. One of the methods to indicate the classification accuracy is an index, which shows the extent to which the classified distribution fits the original distribution. The method most often used for measuring this fit is the ‘Goodness of Variance Fit (GVF)’. This article outlines briefly how the GVF index is calculated, and thereafter examines the implications of this measure for a case example in the two counties South- and North-Trøndelag located in central Norway (Figure 1).

‘Goodness of Variance Fit’

The GVF is a measure for the statistical conformity between the resulting groupings after a classification and the original data values. A larger homogeneity within classes (small dispersion) and simultaneously a larger heterogeneity between classes (large dispersion), gives a better classification after this criteria. The GVF value is calculated as followed [Dent, 1996, p. 136]:

> Calculate the mean for the variable and calculate the sum of the squared deviation of each observation values in the total array from this array mean. This value is called the SDAM (squared deviation of the array mean) (see frame, formula 1).
>
> Compute the class means. Calculate the squared deviations of each observation value within a certain class from its class means. Each class is assigned this value, which is called the SDCMc. Finally; calculate the sum of all SDCMc’s. This value is called the SDCM (Standard Deviation of Class Means) (see frame, formula 2).
>
> Calculate the GVF value as the difference between the SDAM and the SDCM divided by SDAM (see frame, formula 3).

The GVF values will in theory always be between 0 and 1. Generally, the GVF value increases with an increased number of classes. For choropleth maps, where the amount of classes corresponds to the number of units with unique values in the dataset (classless choropleth maps), the GVF value will be equal to 1 since the 'classified' distribution is identical to its original.

Example

Figure 2a shows the amount of the age group above 67 years in the total population for the municipalities in the counties of South- and North-Trøndelag in 1999. The variable is classified into five classes according to the quantile-method which divides the amount of observations equally or nearly equally over the amount of classes. In certain commercial cartographic packages this distribution is offered as the default classification. An old observation seems thus still appropriate:

> [Quantiles] "… have been selected by cartographers wishing to play safe and make sure that some spatial differentiation was portrayed [Evans, 1977:107].

According to Evans, spatial differences are often emphasised when the quantile-method is applied, which is also the case in Figure 2a. A geographical differentiation that is not evidently present in the data set is strengthened in the map, while existing differences are concealed.

Whenever a cartographic classification shows differences that are not, or only to a small extent, present the map 'lies' [Monmonier, 1991]. Lying with maps might be done consciously or unconsciously. To avoid the situation where a user of mapping packages unconsciously is lying with maps, this person should receive a feedback on the correctness of the selected class intervals. Herewith, the GVF value may be helpful. It is also visualised how much each class deviate from the variable's distribution. Figure 2b shows the GVF value and how each class (expressed by SDCMc values) contribute to it. The GVF value is 0,899. Classes 1 and 5 (with SDCMc values 33,9 and 15,72 respectively) contribute mostly to the poor GVF value. Consequently, the internal differences between the municipalities around Trondheim are hidden, while instead, the dramatic contrast between the inland municipalities (the high classes) and the municipalities along the Trondheim fjord (class 1) are shown. Furthermore, according to this classification method several municipalities on the fringe of the county are indicated as having a significant percentage of the aged population, since ten municipalities (instead of two as in Figure 3) are classified in the lowest class.

Use of the GVF value

The theoretical maximum value for GVF is 1,0, which means that classifications are better when they have a value closer to 1,0. For this reason the GVF can be used to optimise the classification [Jenks & Caspall, 1971], and there are algorithms applying it [Jenks, 1977; Lindberg, 1990]. The quest for the optimal map illustrates cartographers' occupation from about 1970. This period is called the communication paradigm within cartography [MacEachren, 1995]. Within this paradigm, the map is regarded as a medium for communicating a certain and known message. After about 1990 cartographers' interests shifted towards a visualisation perspective. According to this perspective, one is no longer engaged in searching for the optimal map:

> For cartographic visualization the message is unknown and, therefore, *there is no optimal map*! The goal is to assist an analyst in *discovering patterns and relationships* in the data [MacEachren & Ganter, 1990: 65].

In my opinion, the quest for optimising a classified map based on the GVF value is still of importance. A better GVF value will lessen the probability that differences or similarities that are present in a dataset are faded down in the map, or that differences or similarities which are not present in the dataset, appear in the map. Therefore, the GVF value is relevant regardless of whether it is regarded from a communication paradigm or from a visualization paradigm. When the map is used as a device for communication it holds that with a better GVF, the probability decreases that the map will "lie". When the map use is as a tool for visualisation, it holds that with a better GVF, the probability increases that a pattern that an analyst eventually discovers in a map also exists in reality. If the map serves as a basis for decision-making, the importance of an accurate representation becomes even clearer. The calculation of GVF values to optimise the classification, or the use of GVF as a remedy for selecting among various classification methods [Declercq, 1995] remains, thus, relevant.

One of two hundred thousand

The map in Figure 2a is just one of 194.580 possible for this variable with 49 units divided into five classes. It is meanwhile difficult to judge whether the classification, used for the map in figure 2a, is good or poor, without knowing the maximum GVF possible for this dataset. Therefore, a Visual Basic algorithm is developed to calculate a GVF value for each of the 194.580 possible classifications, which determines 0,937 as the maximum value. Declercq [1995:922] investigated how the number of classes influences the GVF values for various classification methods, which showed that the GVF values level with values over 0,95, even though the amount of classes increases. From this he concluded that a GVF value of 0,95 might be viewed as the norm for an accurate classification. It appeared that for a classification of a variable with 308 units, six classes were needed to reach this GVF value. Following Declercq's conclusion, there exists no normatively good classification for our dataset when using five classes. However, by increasing the number of classes to six, a GVF value of 0,954 is arrived at when class intervals are determined according to equal intervals. The GVF value will meanwhile only amount to 0,907 with a classification after the quantile method.

When making a map from a dataset, a classification method should be chosen that gives a GVF that is as close as possible to the optimal GVF. In practice, with commercial packages this implies choosing the classification method which gives the best GVF value. For our dataset it appeared – by using five classes – that the equal interval method gives the best GVF value (see Figure 3).

A visual comparison between the graph of the original distribution (Figure 4a) and the ideal form the distribution should resemble to be appropriate for a certain classification method (the theoretical curve, see Figure 4b), could be a useful tool for a user of cartographical packages who needs to choose an appropriate classification method. When the graph equals e.g. a linear function, it is appropriate to apply the equal intervals as classification method [Ormeling & Kraak, 1987:78, Figure 3-30].

Sometimes, the visual interpretation of the form of the original distribution is not unambiguous. In Figure 5 three curves for quantitative series are reproduced: (a) linear, (b) normal and (c) progressive convex distribution of data values. These distributions are successively adequate for (a) equal intervals, (b) standard deviation as dispersion unit, and (c) geometric progression (increasing convex). A visual inspection would conclude that all three curves are more or less comparable with the original distribution (Figure 4a). When such doubt arises, the GVF value can be used to find out which of the methods that best fits the original data. The GVF values for the three classification methods with a grouping in five classes are:

Equal intervals: 0,930
Standard deviation: 0,924

Geometric progression (increasing convex): 0,923

Even though all classifications give high GVF values, the method equal intervals is the preferred one as it gives the highest GVF value.

Visualization of the classification inaccuracy

A single number between 0 and 1 will not necessarily be understandable for the user of a cartographic package. The GVF value should therefore be accompanied with a graphic representation making its meaning explicit. This can, for example, be performed by showing the variable's distribution in a sorted diagram with class borders and class means inserted and where a certain colour expresses how the unit values deviates from the class means as in Figures 2c and 3c. In these figures there are the deviations for each class from the class means painted in red. The areas formed by these deviations within each class corresponds to the sum of deviations over all classes, or the classification error. The relative contribution each class has on the classification error is thus visualised. Note that this is not equal to the SDCMc value (see Figures 2b and 3b), which forms the squared sum of deviations within a class. Nevertheless, it would probably be easier to understand a GVF value when seeing a relation between the size of the represented classification error and the GVF. When the areas corresponding to the classification error decline – and the classification error is thus reduced – the GVF value approaches 1. It will also be probably easier to understand which class contributes most to the classification error because the area formed by the deviation of this class is the largest.

It is not unusual to find commercial packages in which the user easily can manipulate the class borders interactively. In the software package Descartes (Andrienko & Andrienko, 1998) the class borders are shown in a 'number line plot' or a dispersion diagram. Mouse movements can alter class borders. If one could do the same on a sorted diagram, the user could experience how changing the class borders influences the classification's accuracy because the areas formed by the classification error changed.

Conclusion

For classified maps it is evident that the probability that geographical patterns are concealed is larger with maps having a low GVF value than with maps having a high GVF value. It would lead to a better understanding of classification accuracy if a user of packages for thematic map production was offered a GVF value and, in addition, a visualisation of each class's contribution to the classification error. In GIS and cartographic packages, the GVF value can additionally be used to offer as default the classification method giving the best GVF value. Eventually, this can make a contribution to improve the communication of that which is known, or that which is yet to be explored.

Figure captions

Figure 1 – The counties South- and North- Trøndelag in Norway.

Figure 2 – Percentage of population above 67 years on the total for each municipality in South- and North- Trøndelag (1999). Classified into five groups according to the quantile method.

Figure 3 – Percentage of population above 67 years on the total for each municipality in South- and North- Trøndelag (1999). Classified into five groups according to the equal interval method.

Figure 4 – A visual support when deciding which classification method to use. To the left is the original distribution for the data set shown, to the right is the dialog window in where the ideal graph for an equal interval classification.

Figure 5 – Three graphs for quantitative series: a. linear, b. normal, c. geometric progression (convex).

Article VI

Rød, J.K. VI. The selection of class intervals revisited. Submitted to *Transactions of the Institute of British Geographers*.

The selection of class intervals revisited

Jan Ketil Rød

Abstract

The methods for selecting class intervals, which can strongly affect the visual impression given by a map, is currently reduced to a minimum in commercial available mapping packages. By reducing the options available statistical mapping becomes easy, but the probability of producing biased maps is increased. This paper outlines, through presenting results from a mapping package prototype, how simplicity might be achieved with a widened set of methods for selecting class intervals by offering recourses to make one's choice as well as reasonable default options. In order to make the statistical mapping process more transparent, instruments for quality control is developed which assess classification accuracy.

key words: Norway data classification choropleth maps interactivity accuracy transparency

Introduction

In 1977, Transactions of the Institute of British Geographers published the article 'The selection of class intervals' written by Ian S. Evans (Evans, 1977). The selection of class intervals is essential for the preparation of quantitative thematic maps with graded symbolisation like choropleth maps, which are frequently used for cartographic visualisation of statistical data stored in geographical information systems (GIS) or in mapping packages. The range of data values, which most often are collected for geographical areas with fixed boundaries such as counties and municipalities, are classified into discrete categories. Most of the methods for data classification are revisited in this article. Details on each method will not be provided here as these might be found in numerous textbooks and articles. Instead, emphasis is laid on the necessary equipments of the user interface of a choropleth-mapping environment with reasonable default options and useful resources for selecting and judging class intervals as well as assessing the classification accuracy. This is done mainly through presenting results from a mapping package prototype called 'GIB' developed by the author and designed to include features that promote an interactive cartographic visualisation of statistical data. It allows, for instance, the users to select from a number of classification methods in contrast to the more restricted set of methods usually being available. The variable: *Number of places in institutions for the elderly per 1000 inhabitants above 67 years in Trøndelag in 1997* is provided by Statistics Norway and used in GIB in order to produce case examples for this article. Trøndelag is a region in the middle of Norwegian which consisting of two counties: the southern and northern Trøndelag.

The GIB software package has been developed as part of my Ph.D. studies and is used in the first year geography course at the Norwegian University of Science and Technology. Basic cartographic techniques form a central part of the curriculum for this geography course where students learn to use statistical mapping as a device for presenting geographic knowledge and as a tool for exploring geographic data in order to gain new geographic knowledge. Hence, there was a need for a mapping package functioning as a tool for geographic exploration, for which interaction is recognised as paramount (MacEachren and Ganter 1990, 74). The users should be able to do a wide variety of things to their geographic data and to visualise these in numerous ways. As first year students are likely to be unfamiliar with mapping packages, these should as well be relatively easy to use. Several mapping packages are indeed easy to use, but do not allow for the interactivity as required. Related to the selection of class intervals, requirements on interactivity entail a number of choices, which make the mapping situation cumbersome. The issue was thus whether it was possible to develop a mapping package for choropleth mapping combining the ease of use with interactivity. Simplicity is in the GIB package achieved by giving the users resources to make the choices and by offering reasonable default options. An additional and essential aim was to include a way of evaluating the statistical accuracy of the choropleth map (its quantisation error) and to make the students aware of the possibility to use choropleth maps in order to promote particular mapped interpretations. This functionality, which is found in the GIB mapping package, represent a quality control that makes the mapping situation for choropleth maps more transparent. A private company in Trondheim, Norway, has found this to be of commercial interest and will hopefully be able to offer a commercial available product by the end of 2001.

To group or not to group?

Choropleth mapping include here both the production of classed and unclassed choropleth maps. 'The main argument in favor of using class intervals seems to be that their use enhances readability' (Tobler 1973, 264). Tobler called this argument as an assertion and questioned 'why the theory for pictures' (having near-infinite gradations) 'should differ from the theory for choropleth maps, since both have visual information processing as their ultimate objective' (Tobler 1973, 264). Tobler argued that the choropleth map without class intervals, 'on which the visual intensity is exactly proportional to the data intensity' have 'no quantization error' and consequently, the 'difficult problem of optimum class intervals are thus circumvented' (Tobler 1973, 262). Dobson criticized Tobler's solution since it had no satisfactory solution to the 'perceptual error (an increasing function of the number of classes)' (Dobson 1973, 359). Others advocated the use of choropleth maps without class intervals and improved the method (Brassel and Utano 1979; Peterson 1979). Kennedy (1994) revisits the debate over the merits of using classed versus unclassed choropleth maps challenging (if not debunking) the main argument against unclassed choropleth maps: degenerating readability. Kennedy thus challenges the assertion that the high accuracy of unclassed choropleth maps is only mathematical and not perceptual. 'Map readers are able to regionalize spatial patterns on unclassed choropleth maps, despite the large number of classes, and they are able to perform value discrimination quite well on unclassed maps' (Kennedy 1994, 19). Andrienko and Andrienko have solved the difficulties of degenerated readability in classless choropleth maps with a dynamic visual comparison technique. 'With this tool some number *N* within the value range of the shown variable is interactively selected, and the map is immediately redrawn using a *diverging*, or *double-ended, colour scheme*' ... 'Values higher than *N* are encoded by shades of one colour (hue), and those lower than *N* are shown by shades of another colour' (Andrienko and Andrienko 1999, 363 – originally emphasis). The reference value *N* is controlled, for instance, by clicking on an object in the map. In order to figure out, based on a classless choropleth map, the right order of compared objects that are spatially disjoint and have surroundings with different degrees of darkness, 'it is sufficient to click one of them and observe in what hue the other is repainted' (Andrienko and Andrienko 1999, 364).

Kennedy's motivation for revisiting the debate about unclassed choropleth maps seems to be that not all map-makers are trained in cartography making the debate 'even more relevant today because currently available technology makes it possible for anyone with a computer and a laser printer to produce scientifically unsound but aesthetically attractive maps' (Kennedy 1994, 16). Consequently, she maintains that there is 'an obligation to communicate the structure of the underlying data' (Kennedy 1994, 20). Others have forwarded similar expressions in the context of classed choropleth map production, which might be the reason why an analysis of histograms, graphic arrays, number line plots or clinographic curves is presented as essential preceding the selection of class intervals. 'To learn the nature of the data the first step is simply to arrange all the values (one for each data area) in a rank-order listing. Plotting the values this way on a linear scale will show the gaps and the clusters in the distribution of values and may be all that is needed to guide the choice of a suitable classification scheme' (Cuff and Mattson 1982, 38). The distribution of values might take several different forms among different sets of data. An alternative to the classless choropleth map approach, which also will improve the numerical data relation, is to use a classification method that better fits the particular distribution of the numerical values that is to be mapped. This latter approach, which requires a variety of classification options to be available, is emphasised in this article. However, an algorithm generating unclassed choropleth map is implemented in GIB based on Kennedy's recommendations (Kennedy 1994, 24). An unclassed choropleth map is produced if the user selects this option while determining the number of classes. The question of classed versus unclassed choropleth map becomes therefore 'a question of how many classes' (Cromley 1995).

Also Slocum (1999, 75) argues, as Kennedy (1994) does, and suggests to use classless choropleth maps if maintained numerical data relations are intended. Shades on unclassed maps are 'made directly proportional to the values falling in each enumeration unit, thus maintaining the numerical relations among the data' (Slocum 1999, 75). Slocum, however, seems to follow the main argument against unclassed choropleth maps. 'Although unclassed maps do a better job than classed maps of portraying correct data relations, a disadvantage is that for skewed distributions the ordinal relations in much of the data may be hidden' (Slocum 1995, 75). With ordinal relations, Slocum means whether, for instance, municipalities in one region have lower or higher values than municipalities in another region. 'When the data are classed, these differences become more obvious' (Slocum 1995, 75). It is probable, however, that an inaccurate classified map will enhance an impression of geographical difference that only to a small extent, or not at all, is present in the original data values, or that an inaccurate classified map will hide differences that indeed is present in the data. This is a well-known issue among cartographers who have devoted considerable energy to develop optimal data classification, for instance for choropleth maps (Jenks 1977). More recently, the search for an optimal solution seems to have been replaced by an exploratory approach. 'Rather than developing expert systems that help find a single optimal map for representing a set of information, we need to develop systems that encourage exploration of multiple perspectives on the same data' (MacEachren and Ganter 1990, 75). Exploremap (Egbert and Slocum 1992) and Descartes (Andrienko and Andrienko 1999) are examples on the latter approach. While developing the GIB package, the approach of optimising the classification is tried combined with an approach that encourage exploration. The present paper shows how information about the classification accuracy can be used in order to support class interval selection and, subsequently, to evaluate its “goodness”. In both instances, this information indicates the correspondence between the geographic reality and its mapped representation. ‘In many cases we seem to have lost sight of the fact that maps are intended to communicate something about geographic reality. We have instead limited our attention to evaluating the reader’s ability to interpret the mapped representation of that reality. To evaluate ‘communication effectiveness’ of a thematic map, we must first know the underlying accuracy of that map’ (MacEachren 1985, 38). The underlying accuracy of a map will, according to MacEachren, for any quantitative thematic map be a function of four factors: (1) map production procedures, (2) data collection methods, (3) data classification strategies and (4) symbolization techniques (MacEachren 1985, 39). While MacEachren’s paper gives specific attention to accuracy related to choropleth symbolisation (the fourth factor), the present paper is directed to accuracy related to class interval selection (the third factor).

Groups of classification methods

There have been several attempts to classify methods of class intervals selection (i.e. Evans 1977, Robinson et al 1984, Cauvin et al 1987). Among them, Evans (1977) has been regarded as the most effective (Coulson 1987, 19). Evans (1977, 100-102) grouped classification methods into four groups:

1. Arbitrary classification schemes where the intervals are chosen without any clear aim in mind.
2. Exogenous classification schemes, that apply external class intervals, determined a priori without regard for the data distribution itself.
3. Idiographic classification schemes, with an interior logic that determines the class intervals with respect to specific aspects of the data set.
4. Serial classification schemes, which create class limits in a direct mathematical relationship to one another.

Examples of serial, idiographic and exogenous/arbitrary classification methods (and whether they are implemented or not in popular mapping packages are provided in Table I. Two widely used GIS packages (ArcView and MapInfo), the cartographic software package

MapViewer, two Norwegian packages for educational use (Statistisk Sett and NSD Stat) and the Excel mapping module have been investigated.

	ArcView 3.2	MapInfo 6.0	MapViewer 3.0	Statistisk Sett 1.0	NSD Stat 1.1.1	MS Excel 2000
Exogenous / arbitrary						
User defined	x	x	x	x	x	
Idiographic						
Quantiles	x	x	x	x	x	x
Equal area	x					
Nested means						
Maximum breaks						
Natural breaks	x	x				
Serial						
Equal intervals	x	x	x	x	x	x
Standard deviation	x	x				
Arithmetic progression						
Geometric progression						

Table I Common classification methods and whether or not they are implemented in some GIS and mapping software. The number of methods for data classification available decline for reduced sophistication. Several methods are not implemented in any of the listed packages although they are well documented in textbooks and articles.

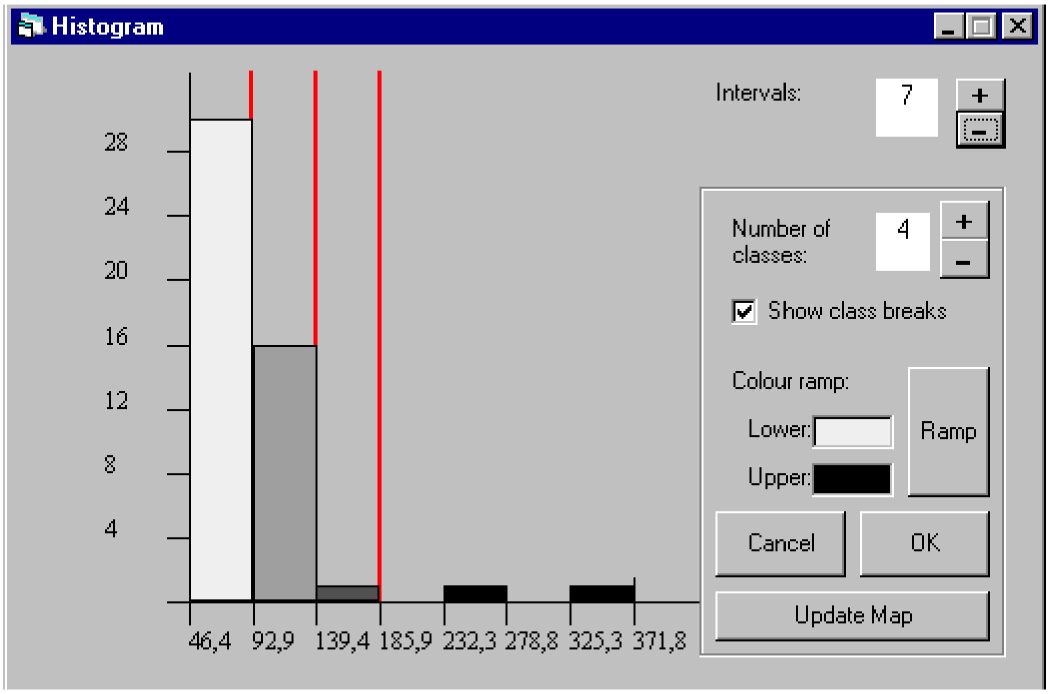
Exogenous classification systems

The user defined classification option often offered in GIS or mapping packages confirms to both an exogenous classification scheme and an arbitrary classification scheme. The class breaks might, for instance, be set according to an intention to identify geographical units (for instance in time series) lying above or below certain values. If no intention is predefined, the classification scheme conforms to an arbitrary one. ‘Arbitrary limits forming no consistent series are indefensible and should never be used’ (Evans 1977, 102). There are instances where exogenous classification schemes might be justified, as for example, in a time series where census classes from different periods need to be matched. These maps are then comparable ‘because given graphic values can be associated with the same numerical values in all the maps’ (Cromley 1995, 16).

Idiographic classification systems

Idiographic class intervals have in common that they are ‘affected by specific details of the data set mapped’ (Evans 1977, 101). Among the idiographic classification schemes listed in Table I, are quantiles, equal area, nested means, maximum breaks and natural breaks. Evans did not favour these classification method, but concluded that ‘idiographic boundaries should almost never be used (Evans 1977, 123).

Evans’ idiographic quantiles (or percentile as he termed it) is ‘classes which contain equal numbers of spatial divisions, *or* near equal areas’ (Evans 1977, 101 – originally emphasis). In Table I, the former of these is called ‘equal area’ while the latter is called ‘quantiles’. A disadvantage of quantiles and equal area is that the methods fail to consider the distribution of the data. The maximum breaks method pays attention only to the largest breaks and thus seems to miss natural clustering of data. ‘The natural breaks method is one solution to the failure of maximum breaks to consider natural groupings of data’ (Slocum 1999, 70). In ArcView, where this method is offered as default, it is implemented as a variant of Jenks’ optimal data classification (Jenks 1977). Otherwise, when this method is considered, it is presented as an interactive, visual process including the examination of a graph or a histogram in order to determine significant breaks. The procedure for constructing histograms involves grouping the data into a number of intervals and it is important to realise that the choice of interval size may have a great effect on the ‘natural breaks’ obtained. Consequently, Jenks and Coulson experienced that ‘subjective judgements based on frequency graphs vary greatly from cartographer to cartographer’ (Jenks and Coulson 1963, 125). In the GIB package, when users select class intervals according to natural breaks, they do so on a histogram as shown in Figure 1.



Intervals: 7

Number of classes: 4

☑ Show class breaks

Figure 1 A histogram of the variable: Number of places in institution for the elderly per 1000 inhabitants above 67 years in 1997 in Trøndelag. The histogram functions as a tool for selecting 'natural breaks'

Since employing differently sized intervals, which is equivalent with employing a different number of intervals, users might change this and realise that the 'natural breaks' also change. Class intervals are automatically set according to the maximum breaks in the histogram and are shown in the histogram as vertical lines. The vertical lines correspond with what Andrienko and Andrienko called delimiters (1999, 362) and can be moved using the mouse. In order to distinguish this type of classification from an arbitrary classification scheme, the rationale behind the class intervals should be illustrated by showing them on a graph or histogram accompanying the legend.

Both the quantiles and the equal area methods apply a principle of equality among classes. Scripter (1970) proposed another approach to equality among classes: the nested means method. The objective of the nested means classification method, 'is to identify the mean as an element of statistical analyses with the goal of creating map classes which maintain the equilibrium system of deviation scores about the mean while generalizing data for visual presentation on a map' (Scripter 1970, 389). The nested means method approximates the equal interval for linear distributions, it approximates the standard deviation method for normal distributions, and it approximates geometric distributions for skewed distributions. Consequently, 'nested means provide the most robust, generally applicable, replicable, yet inflexible class interval system (Evans 1977, 104). The obvious disadvantage of this method – its inflexibility – is that 'the number of classes possible corresponds to the positive integer powers of the number two. For example: 2, 4, 8, 16, 32, 64 … 2^n are the possibilities' (Scripter 1970, 392).

Serial classification systems

Evans recommended the use of serial class intervals (Evans 1977, 98) and perhaps he did so because they have 'limits in a definite mathematical relation to each other' (Evans 1977, 101). The mathematical relation might be illustrated by a curve and some of the more common types of curves include straight line, normal curve, arithmetic progression and geometric progression (see Figure 2).

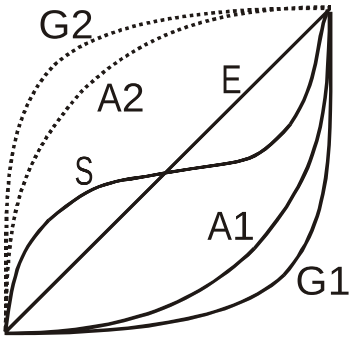


Figure 2 A family of six curves, which may be used to select class intervals. The curves represent (E) straight line, (S) normal curve, (A1 and A2) convex and concave arithmetic progressions and (G1 and G2) convex and concave geometric progressions

Serial classification methods are best applied when aiming at fitting the data classification to a certain function. 'Different frequency distributions suggest different class interval system' (Evans 1977, 102). Hence, the common functions illustrated in Figure 2, each correspond with a particular serial data classification method as listed in Table II.

Common functions	Corresponding methods
Linear (E)	Equal intervals
Normal (S)	Standard deviation
Concave arithmetic progression (A1)	Arithmetic progression
Concave geometric progression (G1)	Geometric progression
Convex arithmetic progression (A2)	Arithmetic progression
Convex geometric progression (G2)	Geometric progression

Table II Correspondences between shapes of distribution and serial classification methods. Text in brackets refers to the common functions illustrated in Figure 2.

Among the serial classification schemes listed in Table I are equal intervals, standard deviation, and arithmetic and geometric progression. The equal interval scheme is the method implemented most. This method was often favoured also before mapping software became available because of its easiness in calculating intervals (Slocum 1999, 67). The equal intervals method is well suited for variables whose values are distributed linearly (see Figure 2).

‘For frequency distributions which are approximately normal, or fairly symmetrical with a pronounced mode near the mean, the standard-deviation system is best’ (Evans 1977, 104). A difficult question, however, is ‘how does one decide whether or not an empirical distribution is normally distributed?’ (Norcliffe 1977, 64). Perhaps it is the relative difficulty of this question that is the reason why the standard deviation method, among the software packages investigated in Table I, is implemented only in the GIS packages ArcView and MapInfo. The practice of using the normal curve as a basis of comparison for a particular distribution is conventional and probably well understood by the more specialised users who are able to investigate the variable’s distributional character in, for instance, a histogram. Such a visual test for normality might be easy for them to perform, but difficult for others that does not have this specific knowledge. The best method to test for normality might therefore be a formal test on the amount of skewness and kurtosis present. Skewness (´$_3$) and kurtosis (´$_4$), also called the third and the fourth moment about the mean, are given by the general formula for the r^{th} moment about the mean:

$$\phi_r \mid \frac{\text{O}(X_i \; 4 \; \overline{X})^r}{N} \qquad (1)$$

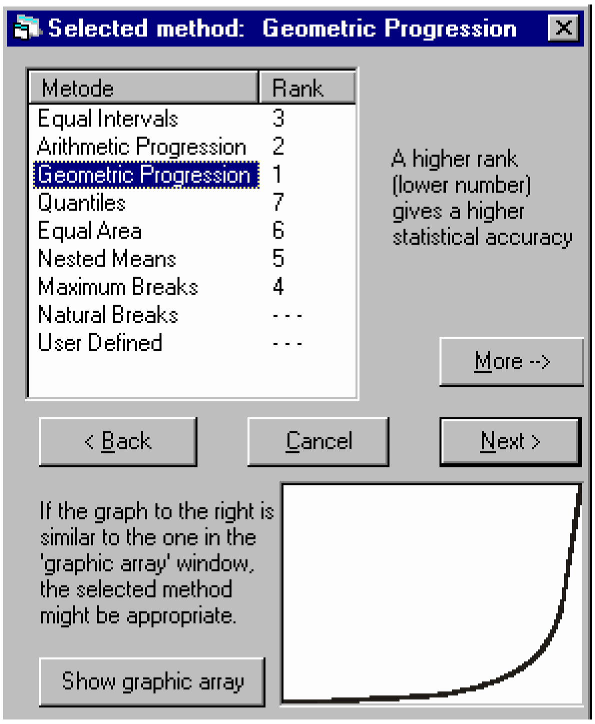
Where N is the number of observations, X_i is the individual data values, and $\overline{X}$ is the arithmetic mean. Skewness describes the extent to which a set of values is slanted in one direction or the other about the mean. Zero skewness indicates a perfectly symmetrical distribution, which will be the case for a perfectly normal distribution. If the value of the skewness increases either by a negative or positive sign, it indicates that the distribution is worse fitted for a method that requires a normal or near normal distribution. Kurtosis is another statistical measurement that can identify whether the data are arranged through a wide range or are peaked. A kurtosis value at approximately three is typical for a perfect normal distribution. In formal tests for normality the values for skewness and kurtosis are converted to relative values known as *ß₁* (beta one) and *ß₂* (beta two) (Norcliffe 1977, 51-52):

$$\eta_1 \mid \frac{\phi_3^{\ 2}}{\phi_2^{\ 3}} \qquad (2)$$

$$\eta_2 \mid \frac{\phi_4}{\phi_2^{\ 2}} \qquad (3)$$

‘In the case of a normal distribution, $ß_1 = 0$ and $ß_2 = 3$’ (Norcliffe 1977, 66). Based on Norcliffe’s modifications of Pearson’s table for the confidence limits for beta one and beta two (Norcliffe 1977, 259), the standard deviation method could be judged to be applicable or not. Such a formal test is implemented in GIB determining whether or not the standard deviation method is made available in the dialog window for selecting data classification method (see Figure 3a).

a)



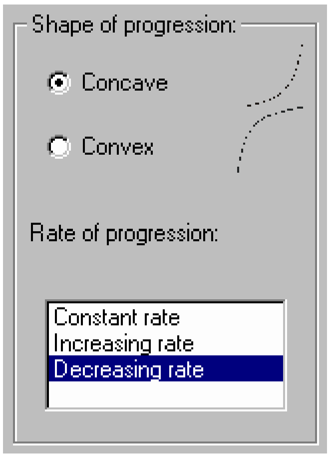
Selected method: Geometric Progression

Metode	Rank
Equal Intervals	3
Arithmetic Progression	2
Geometric Progression	1
Quantiles	7
Equal Area	6
Nested Means	5
Maximum Breaks	4
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

If the graph to the right is similar to the one in the 'graphic array' window, the selected method might be appropriate.

b)



Shape of progression:

● Concave

○ Convex

Rate of progression:

Constant rate
Increasing rate
Decreasing rate

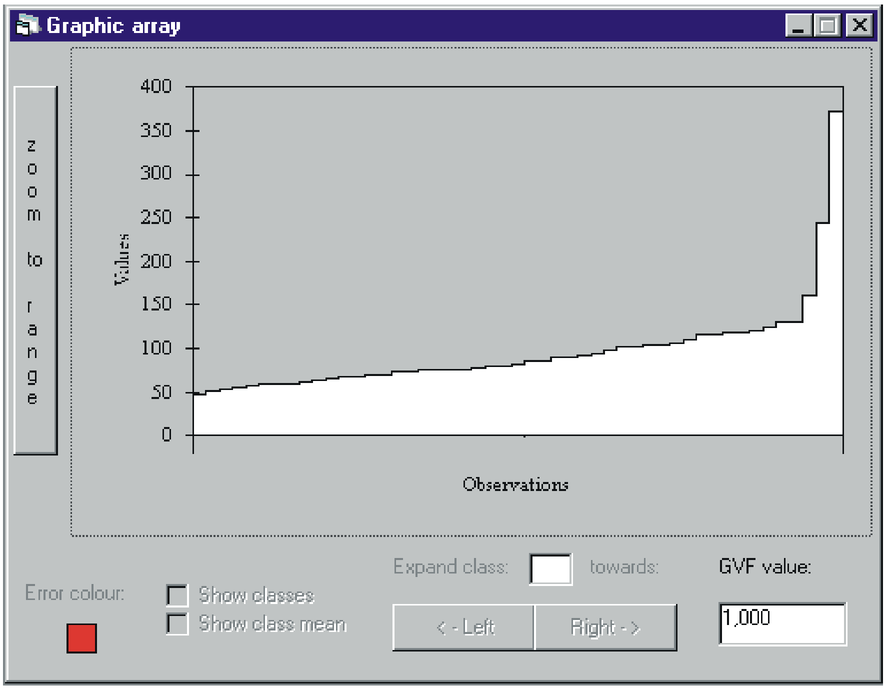
Figure 3 (a) The dialog window for selecting data classification method in GIB. The method with the highest rank is selected as default. (b) For arithmetic and geometric progressions, the 'more' button is visible which, when pressed, will invoke further options for these methods (the one, which gives the most accurate classification, is provided as default).

Although they are well documented in textbooks (i.e. Robinson et al 1995; Dent 1996; Kraak and Ormeling 1996), neither arithmetic nor geometric progression is implemented in any of the mapping packages presented in Table I. These series are flexible as they leave several free choices to the mapmaker who might not know how to make precise decisions. There is, obviously, an infinity of possibilities using arithmetic and geometric progressions. The difficulty is in assessing 'when it is appropriate to apply which particular progression' (Robinson et al 1984, 359). Decisions have to be made regarding the fact whether the progression shall be concave or convex and whether it shall be used with a constant, increasing, or decreasing rate (see Figure 3b). This difficulty might be the reason why arithmetic and geometric progressions are not commonly available in mapping packages. Both methods are implemented in GIB, where, if one of these methods is selected, the combination of options giving the best fit is offered as default.

Simplicity without reducing functionality

There is a general tendency towards reducing functionality for software issued towards a broader public with smaller knowledge and training expected. The more these software packages are sophisticated, the more the user interfaces become complex and the manuals bigger. This translates directly to higher cost for learning and using the system (Frank 1993, 12). Consequently, the interest for user interface design is to achieve simplicity at the expense of functionality: the less options available, the simpler the interface. 'Only when the user interfaces became simpler did wider usage become practical' (Frank 1993, 13). Table I suggests that developers of GIS and cartographic packages follow the approach for user interface design identified by Frank (1993). Simplicity seems to be a prerequisite for a wider use of statistical mapping, which might be the reason why the number of methods for data classification available is reduced for less sophisticated mapping packages. However, as the selection of class intervals strongly can affect the visual impression given by a map, such an approach to simplicity supports what has been called the 'one-map-solution' (Monmonier

1991, 3) and will consequently not be advocated here. This is in particular true for the mapping module in MS Excel where the class breaks are determined according to the equal intervals or quantiles methods only. Making a choropleth map in MS Excel is indeed simple, but also frustrating if the users need to manipulate the class breaks, which they are not allowed. This is indeed autocratic, observing that for a specific number of classes, only two out of several possible choroplethic portrayals is offered. ‘In a choropleth map with N unique data values and p classes, there are (N-1)!/(N-p)!(p-1)! different classification schemes’ (Cromley 1995, 16). It is likely that offering only “two-maps-solution” will enhance an impression of geographical differences that only to a small extent are present in the data, or that it will hide differences that indeed are present in the data. To use Bertin’s words, it is likely that the map will provide ‘false information’ (Bertin 1983, 70).



☐ Show classes
☐ Show class mean

GVF value: 1,000

Figure 4 A graphic array of the originally data sorted by increasing value of the number of places in institution for the elderly per 1000 inhabitants above 67 years in 1997 in Trøndelag

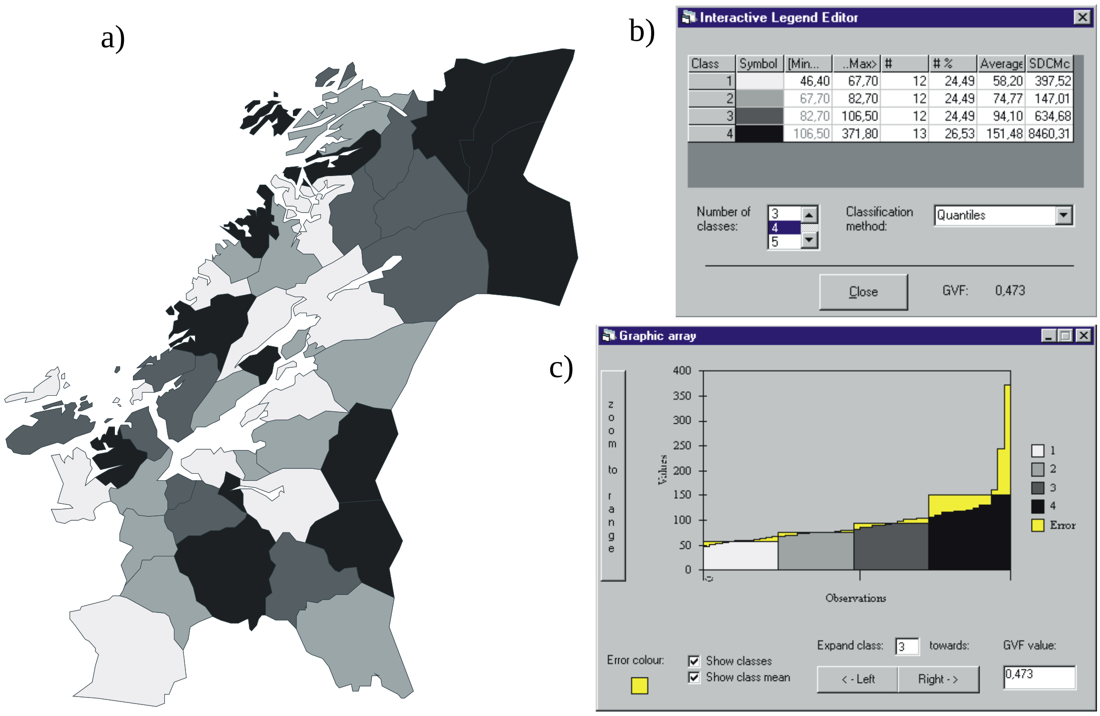
The GIB software package has been developed according to a philosophy of not excluding data classification methods and thus not limiting the number of possible cartographic representations for a given variable. All methods for data classification listed in Table I is therefore implemented. Still, simplicity might be attained by offering reasonable default options and by equipping the interface with simple written and visual responses. In total, GIB provides 10 classification techniques, the user defined option included. For the case variable used in this study, by plotting the variable values against the Y-axis and the individual municipalities against the X-axis, a graphic array showing the shape of distribution appears in another window (see Figure 4).

As seen in Figure 4, the variable is highly skewed. Based on the formal test for normality described above, the standard deviation method is found to be inappropriate and thus is not included in the list of the nine available methods for determining class intervals (see Figure 3a) when this particular variable is grouped into four classes. The number of options is, however, still large enough to make the selection of classification method troublesome. Two instruments are therefore developed in order to help the users making a proper decision. For the serial types of classification methods, users have the opportunity of visually comparing the distribution of the variable with the ideal form it should take if a particular method is appropriate. If the graph on the lower right corner of Figure 3a is similar to the graphic array in Figure 4, the selected method might be appropriate. If the user selects other serial methods, the curve shown in Figure 3a updates accordingly to its ideal form as shown in Figure 2. Additionally, an index called *goodness of variance fit* (GVF) is calculated for all methods except for the natural breaks and user defined ones. Higher GVF values correspond with more accurately classified maps. The classification methods are ranked according to their GVF scores. The method having the highest ranking (lowest number) becomes the default option. The method resulting in the most accurate map for this variable, when classified into four classes, is the geometric progression, which gives a GVF value of 0,933 appearing in the graphic array window (See Figure 4).

The two instruments for decision support just outlined, resemble ideas forwarded by Jenks and Coulson (1963) whose class interval test procedure, which involved ‘testing the various sets of classes to see which best fits the data’ (Jenks and Coulson 1963, 128), were applied for the serial classification methods. A similar class interval test conforms, in the GIB package also to idiographic class intervals. While Jenks and Coulson were concerned that the readers would believe that the test procedure was too involved to be practical (Jenks and Coulson 1963, 128), this article suggests from examples taken from the GIB mapping package that a similar test procedure indeed is practical. Unlike the proposal forwarded by Jenks and Coulson (1963, 128), class intervals in GIB are not been adjusted to make readability easier.

Visual and numerical responses

Figure 5 shows the resulting choropleth map for the case variable classified into four classes according to the quantiles method (a) and two available windows giving numerical and visual responses: the interactive legend editor (b) and the graphic array showing classes and class means (c). The three windows are linked and, consequently, they update according to the changes made. In the interactive legend editor (see Figure 5b and 6b) the users may change the number of classes, manually change individual class breaks or the data classification method. In the graphic array (see Figure 5c and 6c), users might expand a selected class to the left (observations are moved from the class below) or to the right (observations are moved from the class above). Traditional methods developed in thematic cartography are, hence, in the GIB package made interactive allowing the users to make multiple representations of the selected variable.



Class	Symbol	[Min...	..Max>	#	# %	Average	SDCMc
1	□	46,40	67,70	12	24,49	58,20	397,52
2	▒	67,70	82,70	12	24,49	74,77	147,01
3	▓	82,70	106,50	12	24,49	94,10	634,68
4	■	106,50	371,80	13	26,53	151,48	8460,31

Number of classes: 4

Classification method: Quantiles

GVF: 0,473

Error colour: ■

☑ Show classes

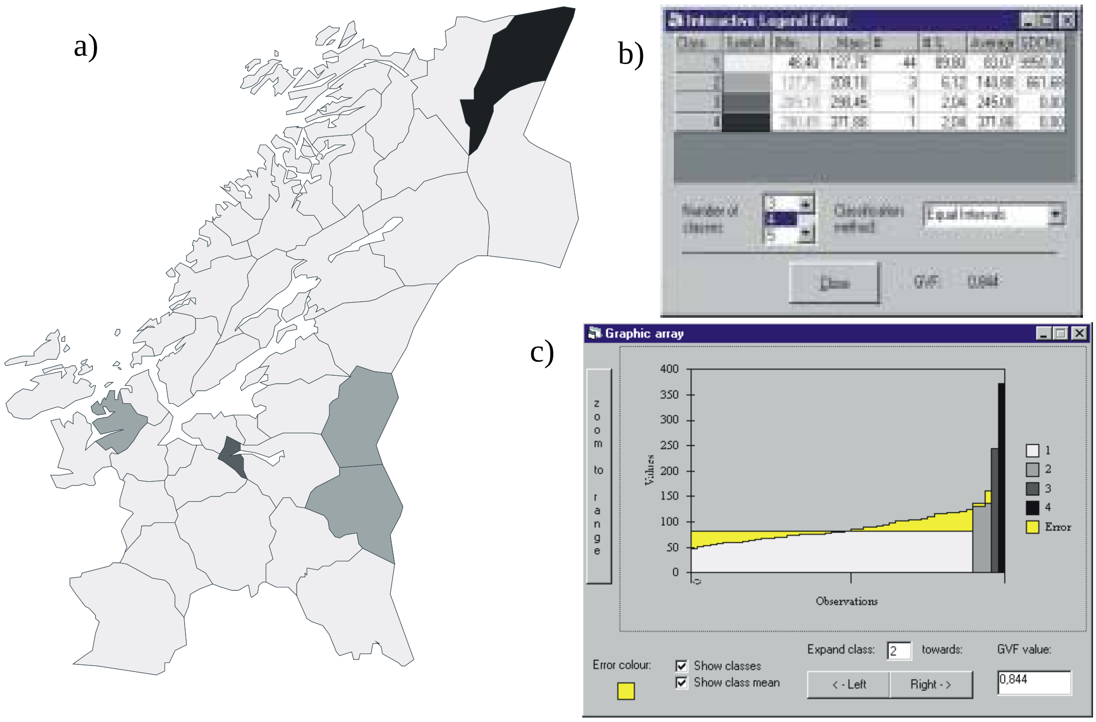
☑ Show class mean

Expand class: 3 towards:

GVF value: 0,473

Figure 5 (a) The number of places in institution for the elderly per 1000 inhabitants above 67 years in 1997 in Trøndelag mapped according to the quantiles method. (b) The SDCMc values indicate each individual class' error contribution. Class four contribute significantly to a poorer fit. (c) Quantisation error is visualised on the graph array showing that the fourth class make the within variation disappear.

The map in Figure 5a shows the number of places in institutions for the elderly per 1000 inhabitants above 67 years in Trøndelag in 1997. The variable with 49 units is divided into four classes with the quantiles method, which divides the observations equally or nearly equally over the number of classes. Commercial mapping packages often use this method as the default data classification method (although with different default number of classes indicated in brackets), MapViewer (6), MapInfo (5), the Excel mapping module (5) and Statistisk Sett (4) included. For this highly skewed variable, most of the municipalities have few institutions for the elderly while a few are in the upper extreme. Applying quantiles as classification method for this variable increases spatial differentiation, but makes it's distributional characteristics disappear. While the graphic array shows that only a few municipalities have high values, several municipalities are classified into the highest class, which strongly affects the visual impression of a high welfare for the elderly because the dark grey tones saturate more map areas. If one closely inspects the graphic array in Figure 5c, one will identify that the fourth class hides significant variance in the original data and that the class breaks between classes three and four are set between nearly equal values and thus emphasising a non-existing difference. The equal intervals map on the other hand (see Figure 6a), seems to visually exaggerate a low welfare level because most of the map area is saturated in a light grey tone.



☑ Show classes

☑ Show class mean

Expand class: 2 towards:

GVF value: 0,844

Figure 6 (a) The number of places in institutions for the elderly per 1000 inhabitants above 67 years in 1997 in Trøndelag mapped according to the equal intervals method in four classes. (b) The SDCMc

values indicate each individual class' error contribution. Class one contribute significantly to a poorer fit. (c) Quantisation error is visualised on the graph array showing that the first class make the within variation disappear

Whenever a cartographic classification fades down significant differences in the underlying numerical distribution, or emphasises insignificant differences, one is lying with maps (Monmonier 1991b). Lying with maps might be done consciously or unconsciously. In order to reduce the amount of unconscious lies, the mapping package users should have some kind of responses informing them about the statistical accuracy of the classification. The interactive legend editor (Figure 5b and 6b) and the graphic array shown with classes, class means and deviations from these (Figure 5c and 6c), is an example on how to design such a response. The class means ($\overline{Z}_c$) composes a generalised model of the individual data values and is indicated by horizontal lines in the graphic array window. Each observation values deviation from its class ($X_i - \overline{Z}_c$) is shown with a specified colour indicated as an error bar (one for each administrative unit). The error bars lie between the individual data values and the class means. By squaring these deviations and summing them for the class, one gets the SDCMc (squared deviations, class mean for each class), which is given in the rightmost column in the interactive legend editor (see Figures 5b and 6b). The SDCMc values indicate the amount of "hidden" variation within each class. By calculating the grand sum one arrives at the SDCM (squared deviations, class mean):

$$SDCM = \sum (X_i - \overline{Z}_c)^2 \qquad (5)$$

SDCM is used in the calculation of the numerical response on the statistical accuracy

$$\mathrm{GVF} = \frac{\mathrm{SDAM} - \mathrm{SDCM}}{\mathrm{SDAM}} \qquad (6)$$

where the SDAM (squared deviations, array mean) is the sum of the squared deviations of each observation (X_i) from the arithmetic mean ($\overline{X}$). The GVF index is a measure of the overall correspondence between the original data values and the class means (see Jenks and Coulson 1963; Jenks and Caspall 1971 and Cromley 1996 for alternative ways to measure error in choropleth maps) and its value will always be between zero and one. The more the GVF value approaches 1.0, the more accurately the generalised model approximates the data model.

How many classes?

High SDCMc values indicate that a particular class contributes to a poor GVF value, like the fourth class in Figure 5 (quantiles classified map) and the first class in Figure 6 (equal intervals classified map). As seen in Figure 5b and 5c, the quantiles method makes the skewness disappear due to the large amount of error contributed by the fourth class. By contrast, the skewness is exaggerated in Figure 6 because of the considerable error contribution from the first class. Still, the latter classification is the preferred since it portrays better the original distribution, which also is reflected by a better GVF value.

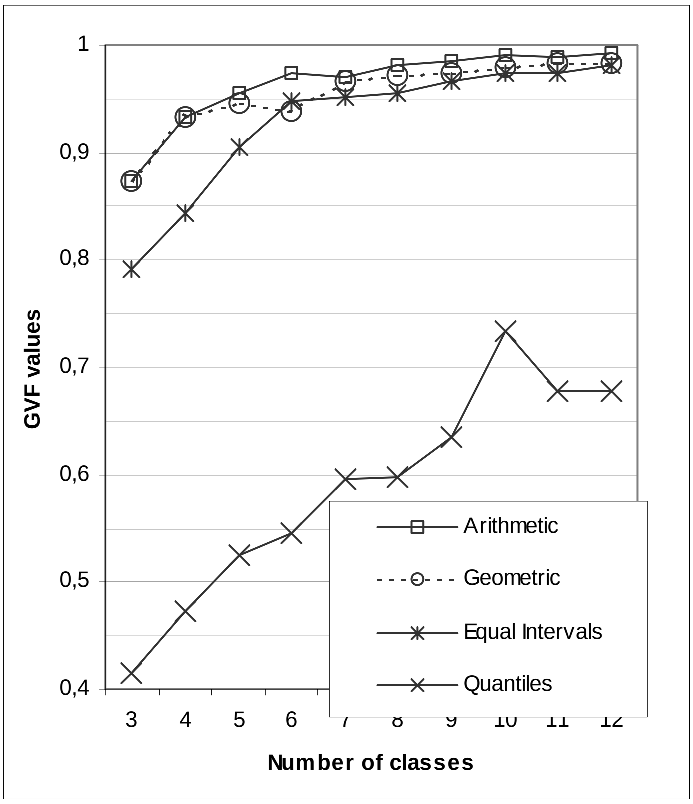


Figure 7 The influence classification method and number of classes have on the GVF values for a classification of the variable *Number of places in institutions for the elderly per 1000 inhabitants above 67 years in 1997 in Trøndelag*

For the variable used in this study, Figure 7 shows a comparison between four different methods for classifying the variable into various numbers of classes. Jenks and Coulson stated that there should be enough classes ‘to avoid sacrificing the accuracy of the data’ (Jenks and Coulson 1963, 120). As commented by Scripter, ‘they did not provide us with a method to determine the number of map classes which will meet this requirement’ (Scripter 1970, 386).

The GVF index, however, could be used for such a prerequisite. As can be seen from Figure 7, generally the GVF value increases with an increased number of classes. For choropleth maps where the number of classes equals the number of units having unique values in the data array (classless choropleth map), the GVF value becomes equal to one as a result of the "classified" grouping being identical with the individual values (consequently, in Figure 4, the GVF value equals 1). In order to judge more generally whether or not a classification method is suited or not for a particular variable divided into a particular number of classes, one needs some criteria. Declercq (1995, 922) considered a threshold value at 0,95 for the GVF value normative for such judgements. Declercq picked the threshold value 0,95 probably because this is about the value were improved GVF values becomes levelled. As seen from Figure 7, for this skewed variable, the most appropriate method is an arithmetic or geometric progression, especially for a small number of classes. For six classes and above, the equal intervals method is approximately at the level for the arithmetic and geometric progressions, while the quantiles method scores poorly also for a high number of classes.

Conclusions

The ruling paradigm in user interface design for data classification seems to be based on a reduced functionality as a mean to obtain simplicity. It is likely that maps produced using the resulting software packages often hide significant differences that appears in the data or make appear insignificant differences. If statistical maps are to be used for studying the geographical distribution of a particular variable, it is essential that the distribution shall be portrayed as accurately as possible. The most accurate choropleth map is the classless choropleth map. Alternatively, more accurate portrayals will also be achieved for classed choropleth maps if the most appropriate data classification method is available and offered as default. The use of numerically quantified and diagrammatically visualised measures of each class' error contribution as well as the overall statistical accuracy (the GVF value), provide a method for quality control more suitable for the selection of a proper number of classes and a proper method for class intervals selection.

References

Andrienko G L and Andrienko N V 1999 Interactive maps for visual data exploration *International Journal Geographic Information Science* 13 (4) 355-374

Bertin J 1983 A new look at cartography in **Taylor D R F** ed *Graphic communication and design in contemporary cartography* Wiley, Chichester 69–86

Brassel K E and Utano J J 1979 Design strategies for continuous-tone area mapping *American Cartographer* 6 (1) 39–50

Cauvin C Reymond H and Serradj A 1987 *Discrétisation et representation cartographique* GIB RECLUS, Montpellier

Coulson M R C 1987 In the matter of class intervals for choropleth maps: with particular reference to the work of George F Jenks *Cartographica* 24 (2) 16–40

Cromley R G 1995 Classed versus unclassed choropleth maps: a question of how many classes *Cartographica* 2 (4) 15–27

— 1996 A comparison of optimal classification strategies for choroplethic displays of spatially aggregated data *International Journal of Geographical Information Systems* 10 (4) 405–424

Cuff D J and Mattson M T 1982 *Thematic maps: their design and production* Methuen, New York

Declercq F A N 1995 Choropleth map accuracy and the number of class intervals in *Cartography crossing borders* 17th International Cartographic Conference Barcelone Sept. 1995 918–922

Dent B D 1996 *Cartography: thematic map design* 4th ed Wm. C. Brown Publishers, Dubuque

Dobson M W 1973 Choropleth maps without class intervals? A comment. *Geographical Analysis* 5 (3) 358–360

Egbert S L and Slocum T A 1992 EXPLOREMAP: an exploration system for choropleth maps *Annals of the Association of American Geographers* 82 (2) 275–288

Evans I S 1977 The selection of class intervals *Transactions of the Institute of British Geographers* New Series 2(1) 98–124

Frank A U 1993 The use of geographic information system: the user interface is the system in **Medyckyj-Scott D and Hearnshaw H M** eds *Human factors in geographical information systems* Belhaven Press, London 3–14

Jenks G F 1977 Optimal data classification for choropleth maps *Department of Geography Occasional Paper 2*, The University of Kansas, Lawrence, Kansas

— and Caspall F C 1971 Error on choroplethic maps: definition, measurement, reduction *Annals of the Association of American Geographers* 61 (2) 217–244

— and Coulson M R C 1963 Class intervals for statistical maps *International Yearbook of Cartography* 3 119–134

Kennedy S 1994 Unclassed choropleth maps revisited / Some guidelines for the construction of unclassed and classed choropleth maps *Cartographica* 31 (1) 16–25

Kraak M J and Ormeling F J 1996 *Cartography: visualization of spatial data.* Longman, Harlow

MacEachren A M 1985 Accuracy of thematic maps / Implications of choropleth symbolization *Cartographica* 22 (1) 38–58

— and Ganter J H 1990 A pattern identification approach to cartographic visualization *Cartographica* 27 (2) 64–81

Monmonier M 1991 Ethics and map design: six strategies for confronting the traditional one-map solution *Cartographic Perspectives* 10 3–8

— 1991b *How to lie with maps* The University of Chicago Press, Chicago

Norcliffe G B 1977 *Inferential statistics for geographers* Hutchinson, London

Peterson M P 1979 An evaluation of unclassed crossed–line choropleth mapping *American Cartographer* 6 (1) 21–37

Robinson A H Sale R D Morrison J L Muehrcke P C 1984 *Elements of cartography* 5th ed Wiley, New York

— **Morrison J L Muehrcke P C Kimerling A J Guptill S C** 1995 *Elements of cartography* 6th ed Wiley, New York

Scripter M W 1970 Nested-means map classes for statistical maps *Annals of the Association of American Geographers* 60 385–393

Slocum T A 1999 *Thematic cartography and visualisation* Prentice Hall

Tobler W R 1973 Choropleth maps without class intervals? *Geographical Analysis* 5 (3) 262–264

Part III

Developed software for statistical mapping

Tutorial I

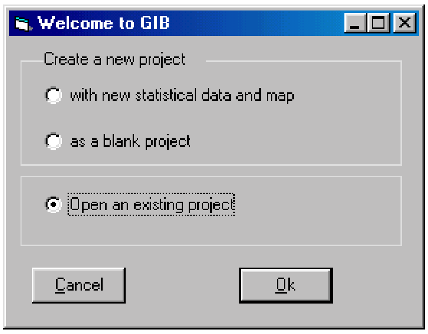
Tutor for the dataset containing article II

TUTOR I

This tutor aims to guide you in using GIB to explore the data set used in *Assessing the functionality of Choropleth Mapping using Excel* (Rød, 1999). You are also guided to explore the dialog boxes for the interactive legend editor shown in the critical comments to this article (Figure II.1) and the dialog boxes for optimal classification shown in the superstructure (Figure 4.4). Each dialog box shown here includes a reference to the GIB manual.

1. Opening an existing project in GIB similar to Figure 1 (Rød, 1999)

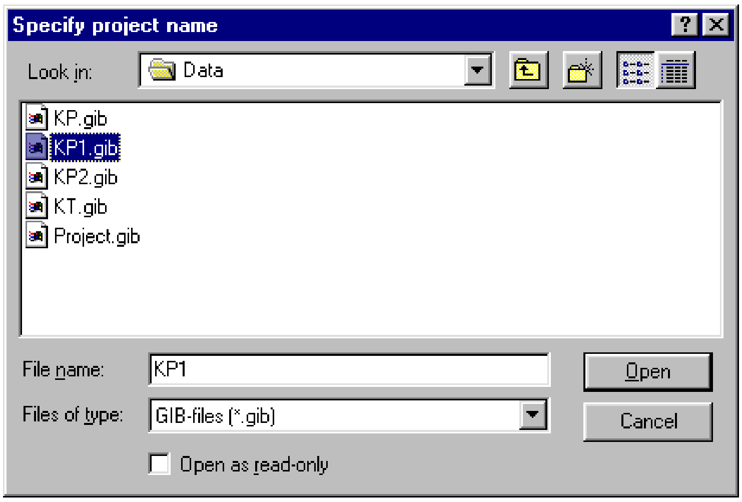
Start GIB
Select *open an existing project*



○ with new statistical data and map
○ as a blank project
● Open an existing project

Ref: Figure 2.1

Select the project *KP1* (abbreviation for *Kart og Plan – Figure 1*).



Look in: Data

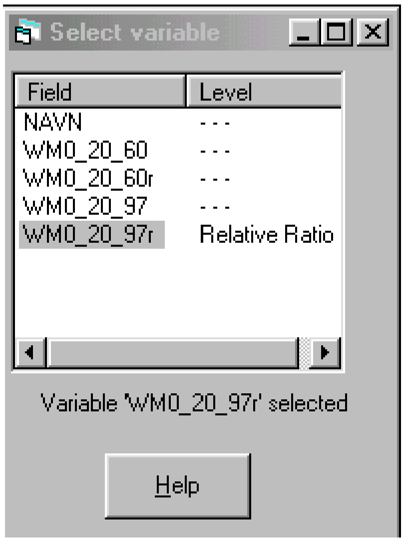
KP.gib
KP1.gib
KP2.gib
KT.gib
Project.gib

File name: KP1
Files of type: GIB-files (*.gib)
☐ Open as read-only

Ref: Figure 2.2

The project will be loaded. For a short moment, an unclassed choropleth map will be displayed before it is replaced by a five-class map for the variable *WM0_20_97r* (Percentage of the population aged 0 to 20 in 1997 in Norway). At your screen, you should have two open windows in the main window, the *Select variable* window and the map window titled “Norway”.

The *Select variable* window should look similar to this:

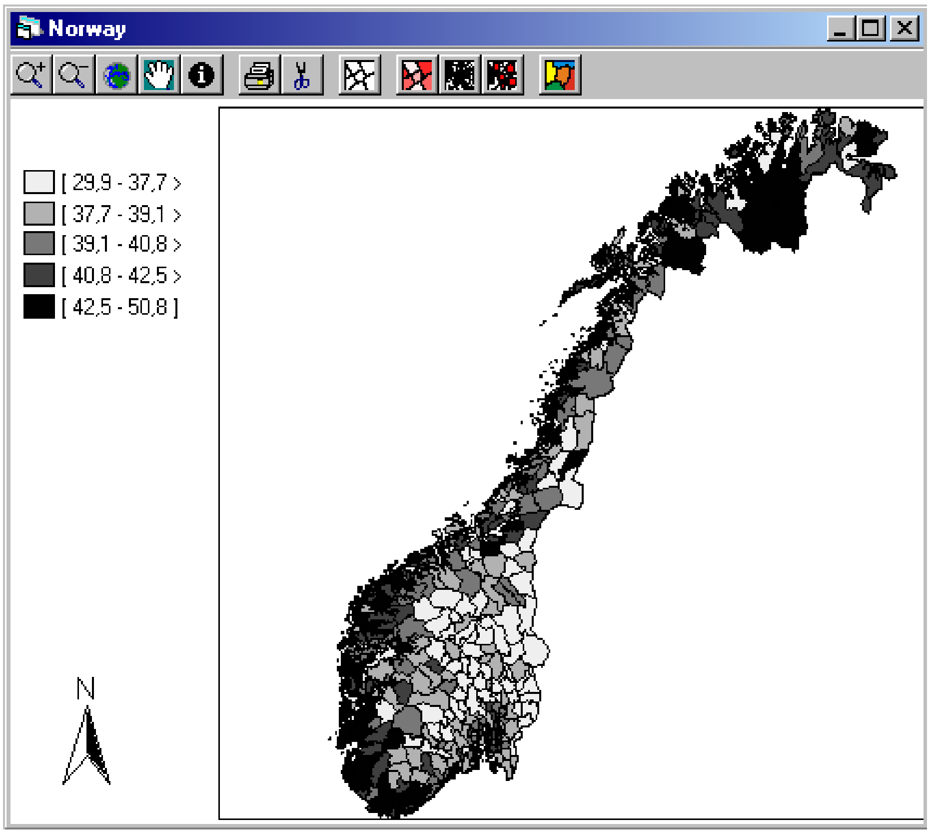


Field	Level
NAVN	- - -
WM0_20_60	- - -
WM0_20_60r	- - -
WM0_20_97	- - -
WM0_20_97r	Relative Ratio

Variable 'WM0_20_97r' selected

Ref: Figure 2.8

The map window should look similar to this:

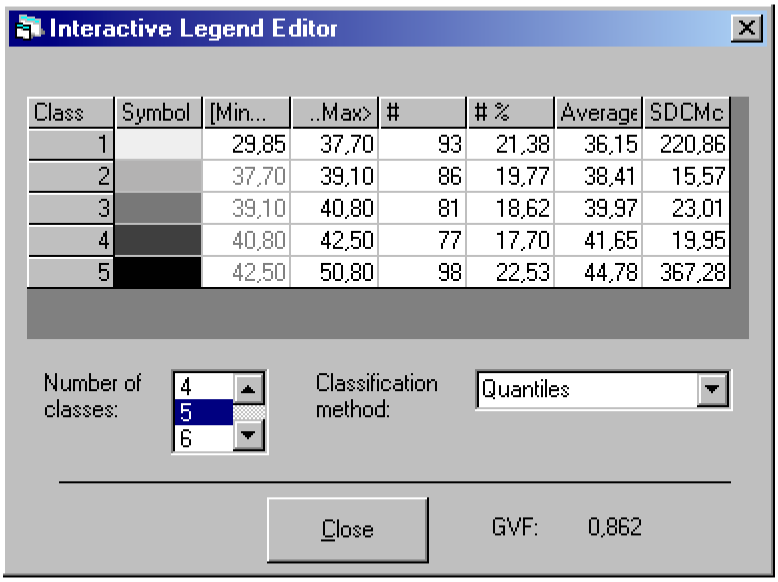


Ref:
Figure 2.7 /
3.11

The project is finished loaded.
Please note that this map is made according to how MS Excel classifies observations into quintiles (five equal or near equal groups). If you reclassify this variable according to the quantile method with five classes, GIB will come up with a different result.

2. Loading the left dialog box from the critical comments (Figure II.1)

While the map still displays, the interactive legend editor can be loaded by selecting *Edit, Edit Choropleth Map*.



Class	Symbol	[Min...	..Max>	#	# %	Average	SDCMc
1	⬜	29,85	37,70	93	21,38	36,15	220,86
2	░	37,70	39,10	86	19,77	38,41	15,57
3	▒	39,10	40,80	81	18,62	39,97	23,01
4	▓	40,80	42,50	77	17,70	41,65	19,95
5	■	42,50	50,80	98	22,53	44,78	367,28

Number of classes: 5

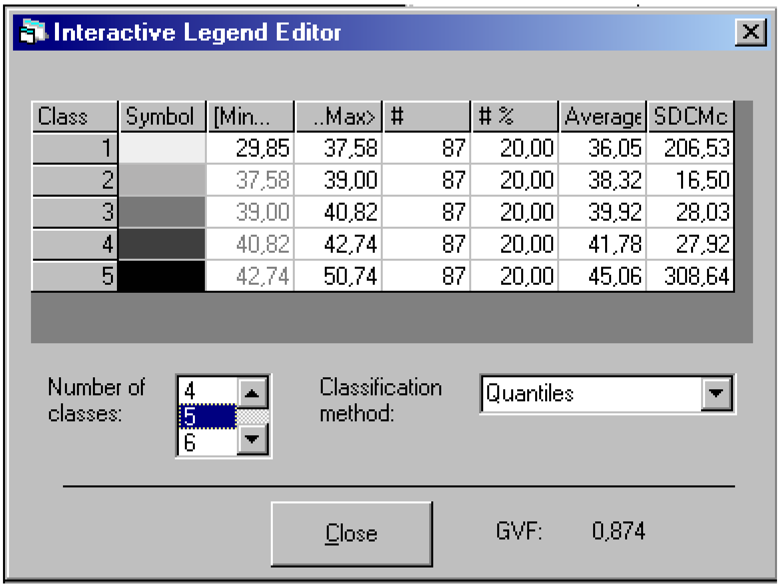
Classification method: Quantiles

GVF: 0,862

The classification scheme with the class breaks obtained using the quantile method in the MS Excel mapping module.

Ref: Figure 4.1.2

Note that this classification scheme equals the MS Excel's quantile division on this variable. GIB will classify the variable differently although using a quantile classification technique.



Class	Symbol	[Min...	..Max>	#	# %	Average	SDCMc
1	⬜	29,85	37,58	87	20,00	36,05	206,53
2	░	37,58	39,00	87	20,00	38,32	16,50
3	▒	39,00	40,82	87	20,00	39,92	28,03
4	▓	40,82	42,74	87	20,00	41,78	27,92
5	■	42,74	50,74	87	20,00	45,06	308,64

Number of classes: 5

Classification method: Quantiles

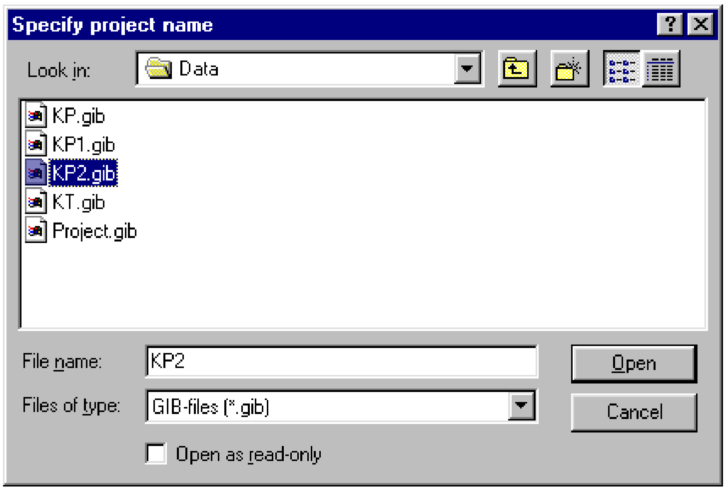
GVF: 0,874

The classification scheme with the class breaks obtained using the quantile method in GIB.

Ref: Figure 4.1.2

3. Opening an existing project in GIB similar to Figure 2 (Rød, 1999)

While GIB is still running, close the active project (select *Close* from the *File* menu) and open a new one (select *Open* from the *File* menu).



Specify project name

Look in: Data

KP.gib
KP1.gib
KP2.gib
KT.gib
Project.gib

File name: KP2

Files of type: GIB-files (*.gib)

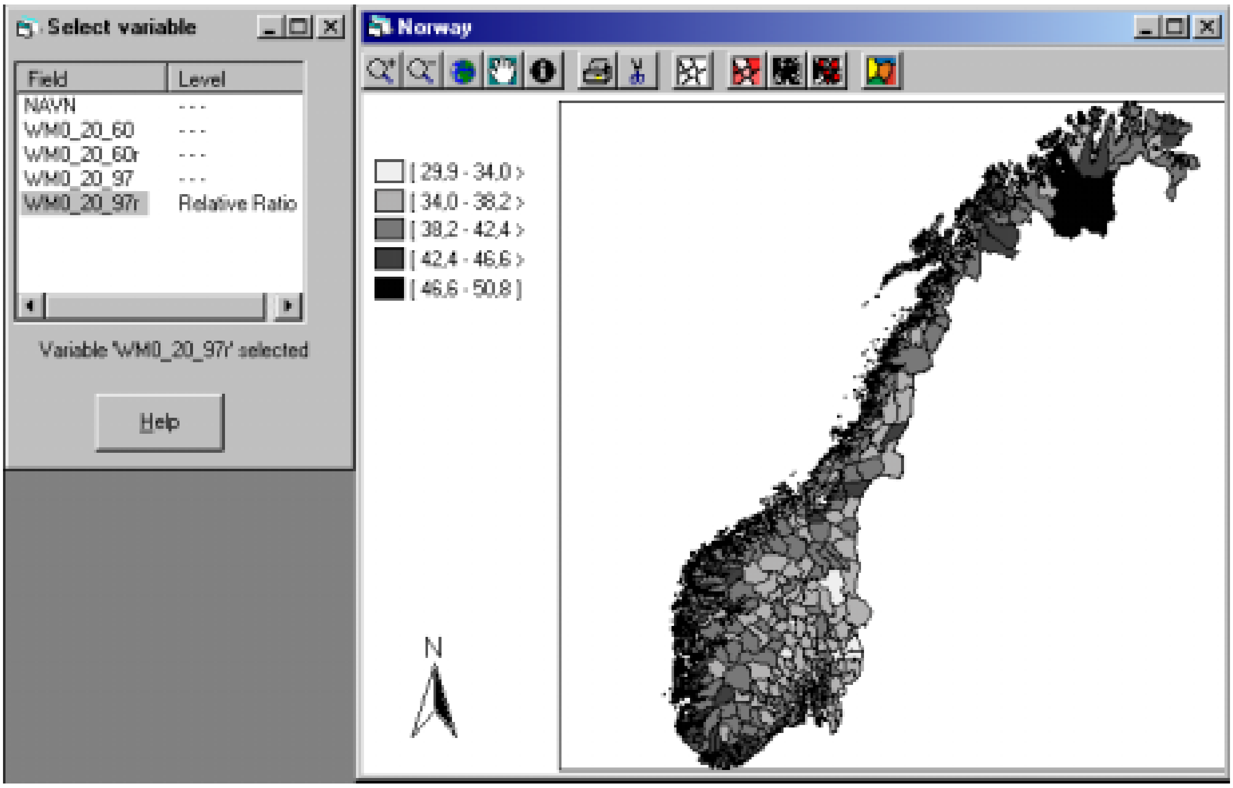
☐ Open as read-only

Ref: Figure 2.2

Select the project *KP2* (abbreviation for *Kart og Plan – Figure 2*).

The project will be loaded. For a short moment, an unclassed choropleth map will be displayed before it is replaced by a five-class map for the variable *WM0_20_97r* (Percentage of the population aged 0 to 20 in 1997 in Norway). At your screen, you should have two open windows in the main window, the *Select variable* window and the map window titled “Norway”.

Your screen should soon look similar to this:



Select variable

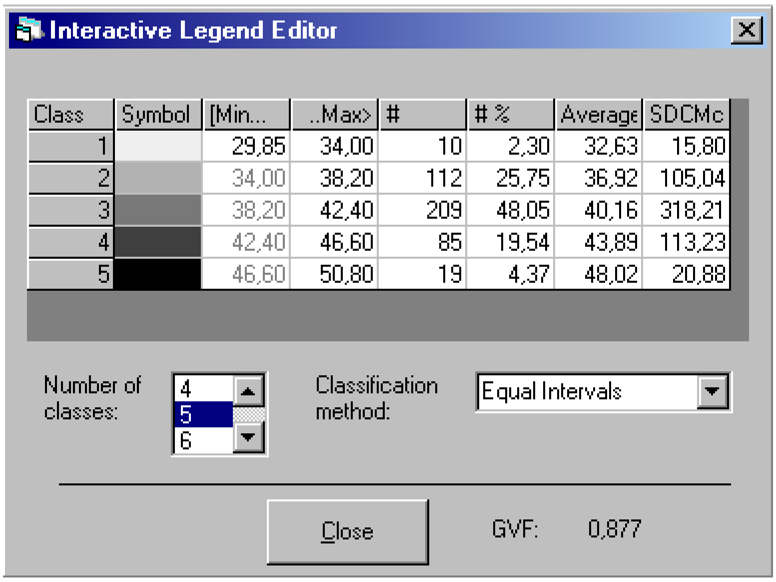
Field	Level
NAVN	- - -
WM0_20_60	- - -
WM0_20_60r	- - -
WM0_20_97	- - -
WM0_20_97r	Relative Ratio

Variable 'WM0_20_97r' selected

The project is finished loaded.

4. Loading the right dialog box from the critical comments (Figure II.1)

While the map still displays, the interactive legend editor can be loaded by selecting *Edit, Edit Choropleth Map*.



Interactive Legend Editor

Class	Symbol	[Min...	..Max>	#	# %	Average	SDCMc
1	░	29,85	34,00	10	2,30	32,63	15,80
2	▒	34,00	38,20	112	25,75	36,92	105,04
3	▓	38,20	42,40	209	48,05	40,16	318,21
4	▓	42,40	46,60	85	19,54	43,89	113,23
5	█	46,60	50,80	19	4,37	48,02	20,88

Number of classes: 5

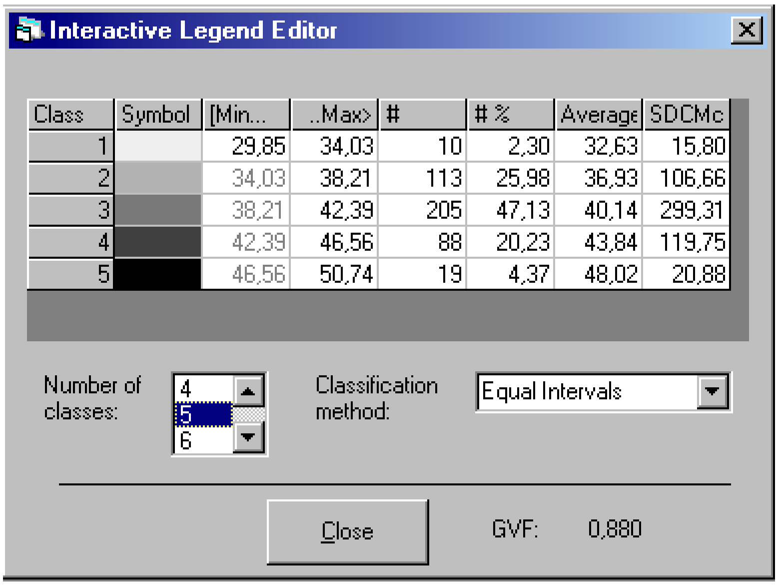
Classification method: Equal Intervals

GVF: 0,877

The classification scheme with the class breaks obtained using the equal interval method in the MS Excel mapping module.

Ref: Figure 4.1.2

Note that this classification scheme equals the MS Excel's equal intervals division on this variable. The difference with how GIB will classify the variable is only slightly different.



Interactive Legend Editor

Class	Symbol	[Min...	..Max>	#	# %	Average	SDCMc
1	░	29,85	34,03	10	2,30	32,63	15,80
2	▒	34,03	38,21	113	25,98	36,93	106,66
3	▓	38,21	42,39	205	47,13	40,14	299,31
4	▓	42,39	46,56	88	20,23	43,84	119,75
5	█	46,56	50,74	19	4,37	48,02	20,88

Number of classes: 5

Classification method: Equal Intervals

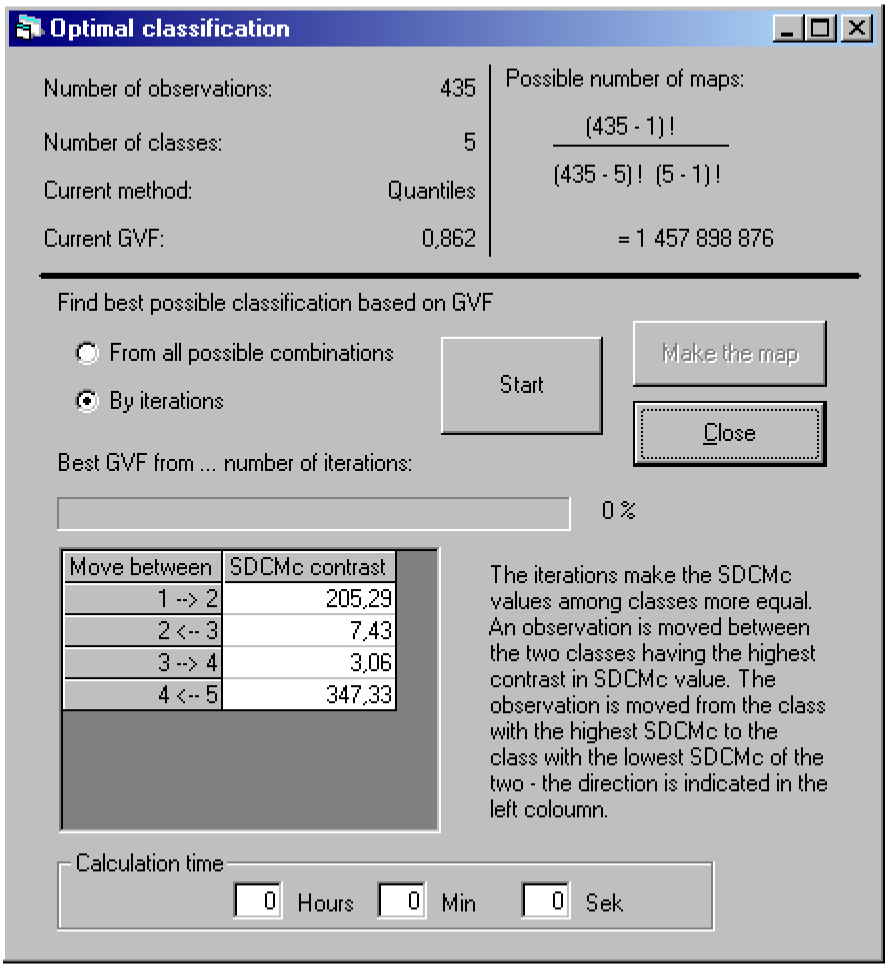
GVF: 0,880

The classification scheme with the class breaks obtained using the equal interval method in GIB.

Ref: Figure 4.1.2

5. Loading optimal classification (Figure 4.4 in superstructure)

While the map still displays, the optimal classification can be loaded by selecting *Edit, Optimal classification*.



Number of observations: 435

Number of classes: 5

Current method: Quantiles

Current GVF: 0,862

Possible number of maps:

$$\frac{(435 - 1)!}{(435 - 5)!\ (5 - 1)!} = 1\ 457\ 898\ 876$$

Find best possible classification based on GVF

○ From all possible combinations

● By iterations

Best GVF from ... number of iterations:

0 %

Move between	SDCMc contrast
1 --> 2	205,29
2 <-- 3	7,43
3 --> 4	3,06
4 <-- 5	347,33

The iterations make the SDCMc values among classes more equal. An observation is moved between the two classes having the highest contrast in SDCMc value. The observation is moved from the class with the highest SDCMc to the class with the lowest SDCMc of the two - the direction is indicated in the left coloumn.

Calculation time

0 Hours 0 Min 0 Sek

Ref:
Figure 4.4.2

The content of this dialog box will depend on the current classification on the map displayed.

For this variable, do not select the option to optimise from all possible combination. The calculations will take very long time.

Click on *Start* to begin the search for an optimal classifcation.
When GIB finishes, the *Make the map* control buttons will be activated. Click on this to update the map.

Tutorial II

Tutor for the dataset containing article V

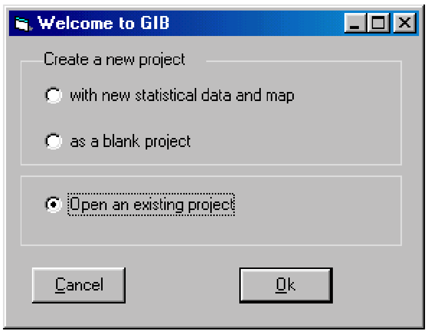
Tutor II

This tutor aims to guide you in using GIB to explore the dataset used in *The accuracy of classified maps* (Rød, 2000). Each dialog box shown here includes a reference to the GIB manual. The data set used in this article is also used to produce Figure 4.2, Figure 4.5 from the superstructure.

1. Opening an existing project in GIB

1.1 If GIB is already running, close the active project (select *Close* from the *File* menu) and open a new one (select *Open* from the *File* menu). Go to 1.3 for further reference.

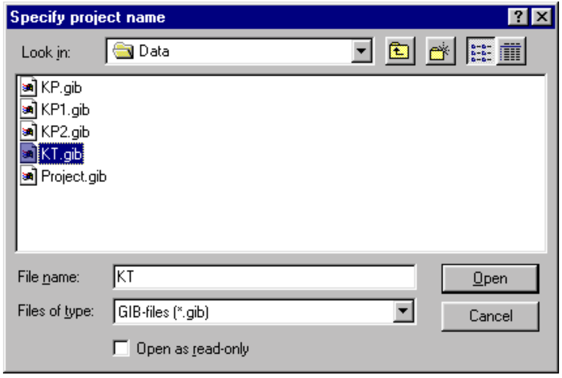
1.2 If GIB is not running, start GIB and select open an existing project



● Open an existing project

Ref: Figure 2.1

1.3 Select the project *KT* (abbreviation for *Kartographisch Tijdschrift*).



Look in: Data

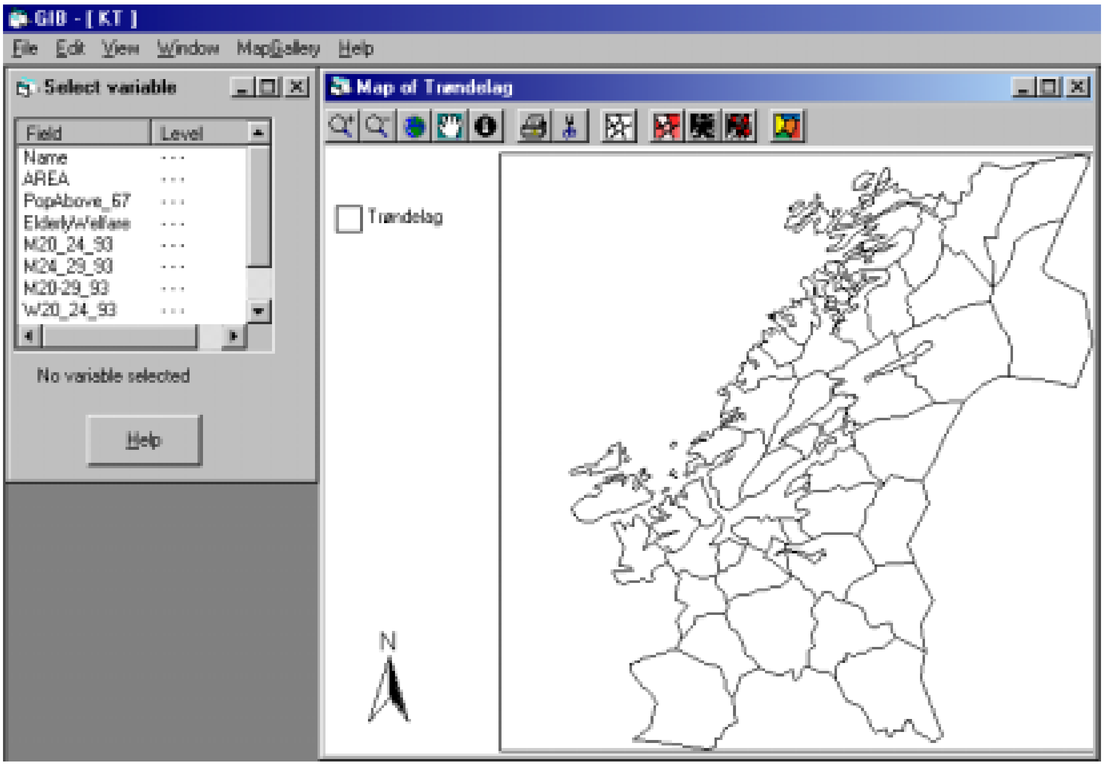
File name: KT

Files of type: GIB-files (*.gib)

Ref: Figure 2.2

The project will be loaded. Please wait until the base map for the two counties “Sør”- and “Nord Trøndelag” is displayed. At your screen, you should have two open windows in the main window, the *Select variable* window and the map window titled “Trøndelag”

For this project, no variable is selected and therefore the message “No variable selected” is shown in the *Select variable* window and the map is shown without symbolisation – as a base map showing only the borders for the municipalities.
Your screen should look similar to this:



The project is finished loaded.

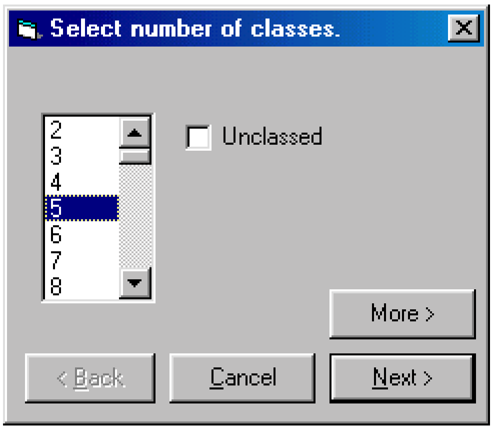
2. The choropleth mapping wizard

Among the variables listed in the *Select variable* window is the variable *PopAbove_67*, which is the percentage of elderly people (men and woman above 67 years) on the total population in 1997 for the two counties. Select this variable by clicking on it.

GIB then investigates the values the selected variable holds and interprets them to be at relative ratio level. A class less choropleth map will then be generated as default.

Start the choropleth wizard (Ref: Figure 3.1.1 / 3.1.2)
- select *Choropleth map* from the *MapGallery* menu, or click on the *choropleth map* icon on the tool menu.

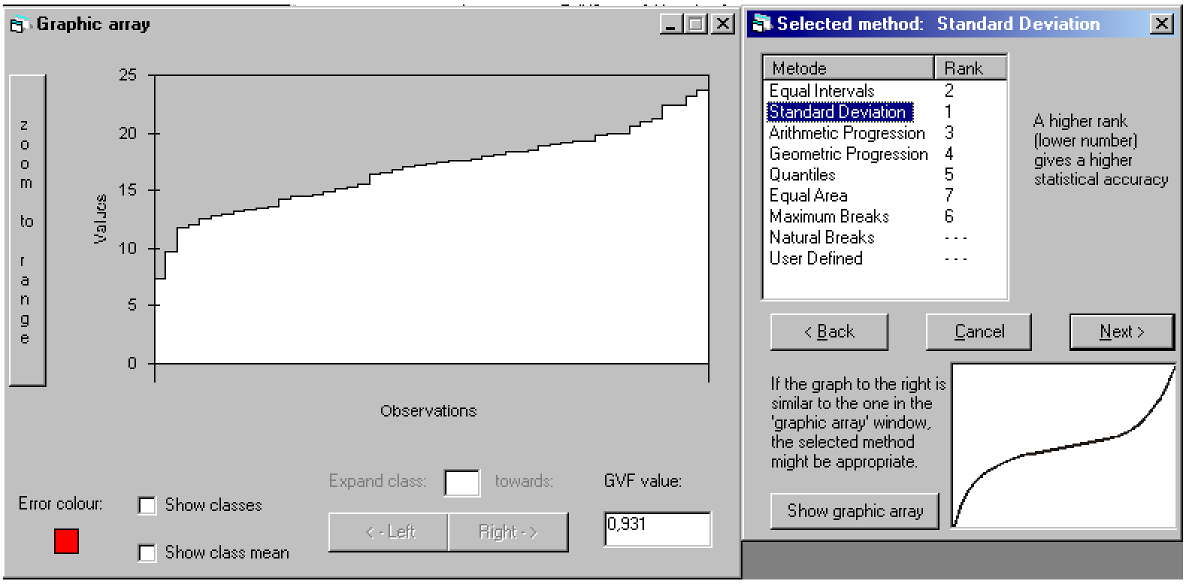
The map will than be regenerated as a base map and the *Select number of class*es dialog box will be loaded:



Select *five* classes
Click next

Ref: Figure 3.2

You will then be displayed the dialog box for selecting classification method, which differ from the dialog box shown in Rød (2000: 50, Figure 4b) with changes outlined in Rød (VI).



GVF value: 0,931

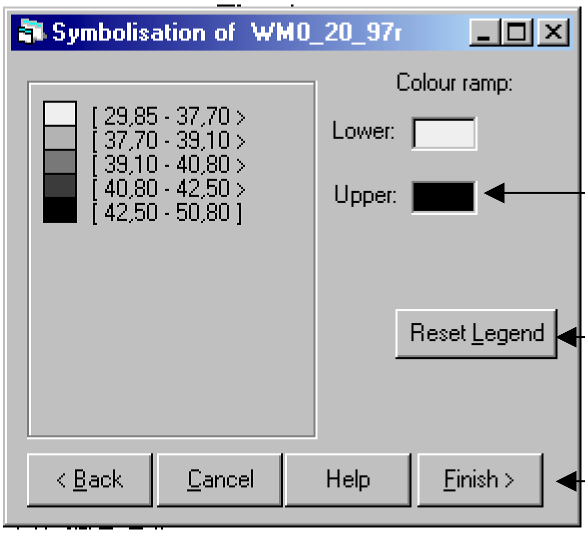
Metode	Rank
Equal Intervals	2
Standard Deviation	1
Arithmetic Progression	3
Geometric Progression	4
Quantiles	5
Equal Area	7
Maximum Breaks	6
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

If the graph to the right is similar to the one in the 'graphic array' window, the selected method might be appropriate.

(Ref: Figures 3.4, 3.5.1 and 3.52)

- Click on the *Show graphic array* button (in Rød, 2000, Figure 4b, this button is called *Show sorted diagram*). A graphic array is then visible in a new window as shown above.
- Click on the *zoom to range* button to achieve a similar diagram to the one in Rød (2000: 50, Figure 4a).
- Click on different method and study how both the "ideal graphs" for the serial methods and the calculated GVF values update.
- Finally, select the classification method *Quantiles*. The dialog box for symbolising the classes is then launched



Ref: Figure 3.9.1

Click on the *Upper* field to change the colour ramp.

Click on the *Reset Legend* to update the legend.

Click on Finish to update the map.

The map will then be symbolised in the colour ramp define

3. Linked windows

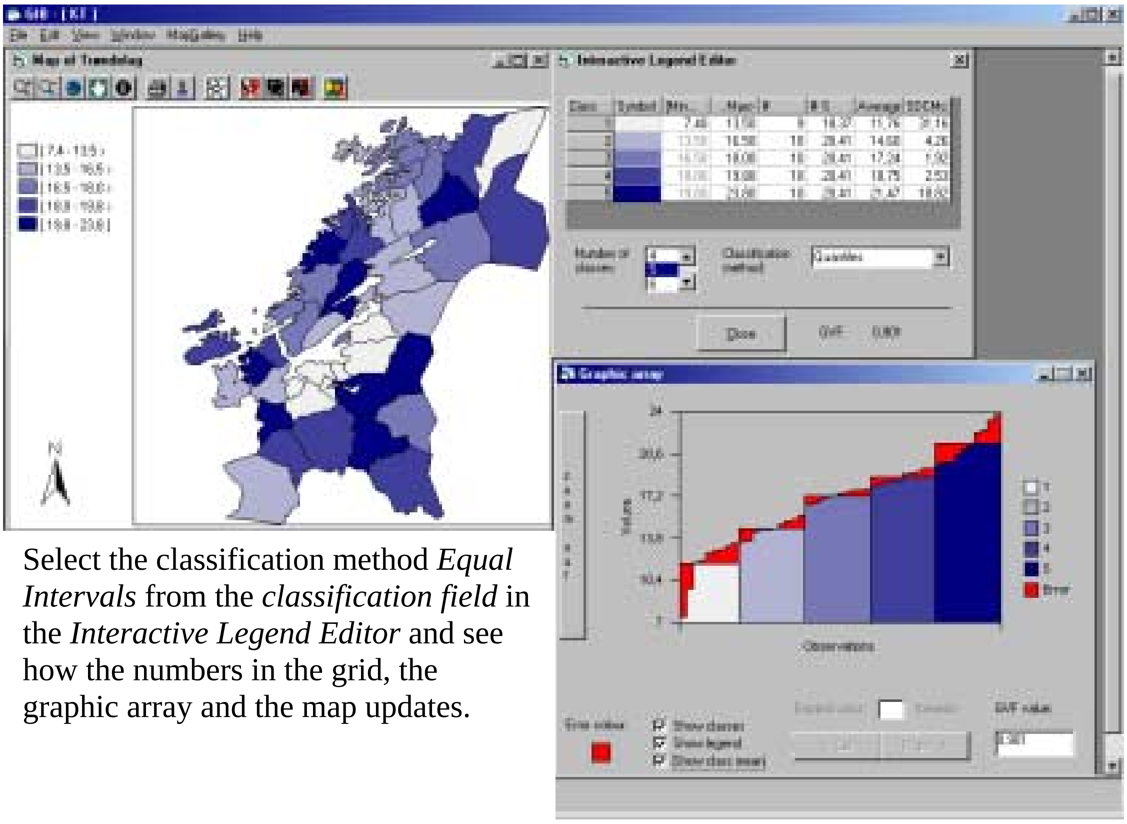
While the map still is displayed, launch the

- Interactive Legend Editor (from the Edit menu, select Edit Choropleth Map)

and the

- Graphic Array (from the *View* menu, select *Graphic Array*)
- activate *show classes* and *show class means*

and situate the dialog boxes on the screen along each sides. For instance something like this:



Select the classification method *Equal Intervals* from the *classification field* in the *Interactive Legend Editor* and see how the numbers in the grid, the graphic array and the map updates.

Tutorial III

Tutor for the dataset containing article VI

Tutor III

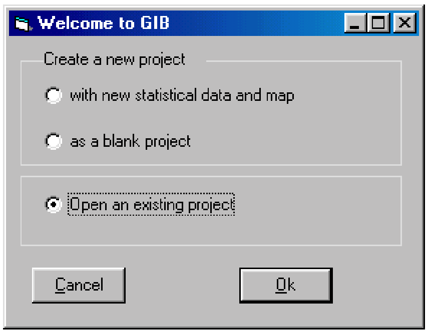
This tutor aims to guide you in using GIB to explore the dataset used in *The selection of class intervals revisited* (Rød, VI). Each dialog box shown here includes a reference to the GIB manual.

1. Opening an existing project in GIB

1.1 In this Tutor you will be using the same project as in Tutor II. If GIB is running and this project is open, go to 2.

1.2 If GIB is already running with another project than *KT*, close the active project (select *Close* from the *File* menu) and open a new one (select *Open* from the *File* menu). Go to 1.4 for further reference.

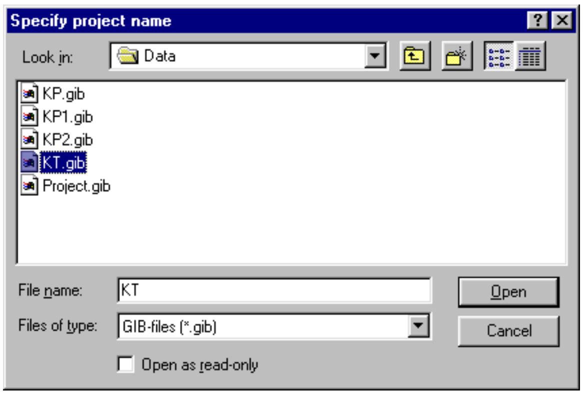
1.3 If GIB is not running, start GIB and select open an existing project



Create a new project
○ with new statistical data and map
○ as a blank project
● Open an existing project

Ref: Figure 2.1

1.4 Select the project *KT* (abbreviation for *Kartographisch Tijdschrift*).



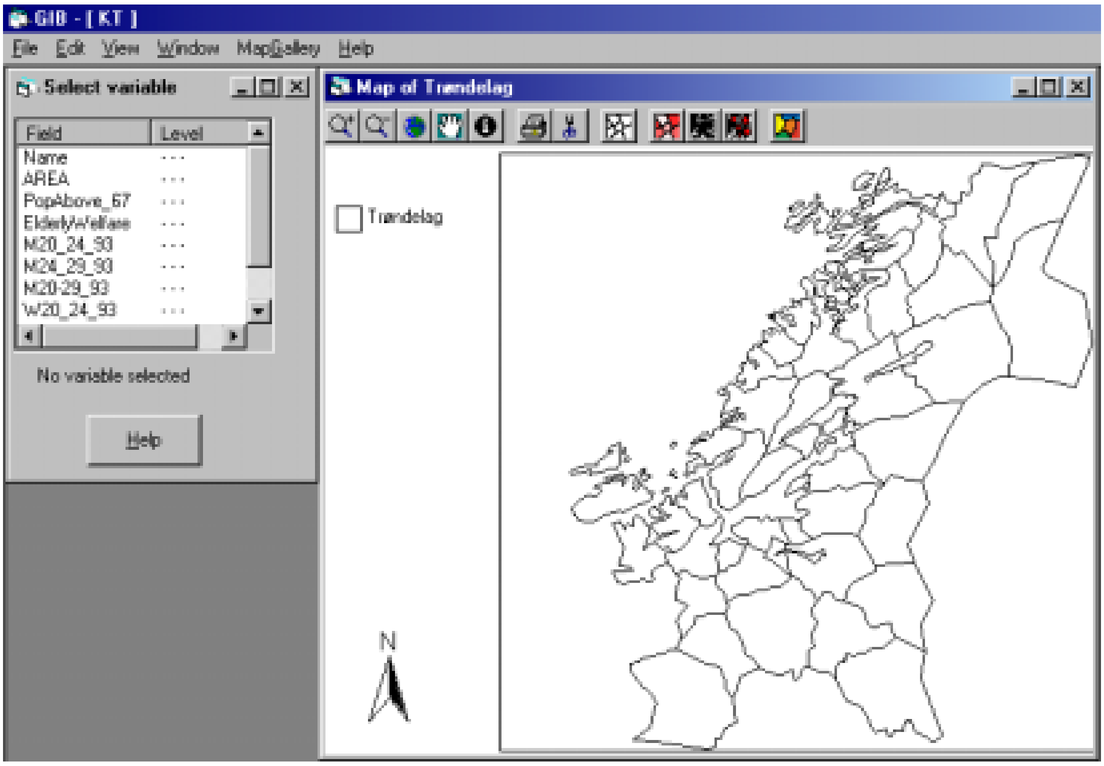
Look in: Data
File name: KT
Files of type: GIB-files (*.gib)
☐ Open as read-only

Ref: Figure 2.2

The project will be loaded. Please wait until the base map for the two counties “Sør”- and “Nord Trøndelag” is displayed. At your screen, you should have two open windows in the main window, the *Select variable* window and the map window titled “Trøndelag”

For this project, no variable is selected and therefore the message “No variable selected” is shown in the *Select variable* window and the map is shown without symbolisation – as a base map showing only the borders for the municipalities.

Your screen should look similar to this:



No variable selected

The project is finished loaded.

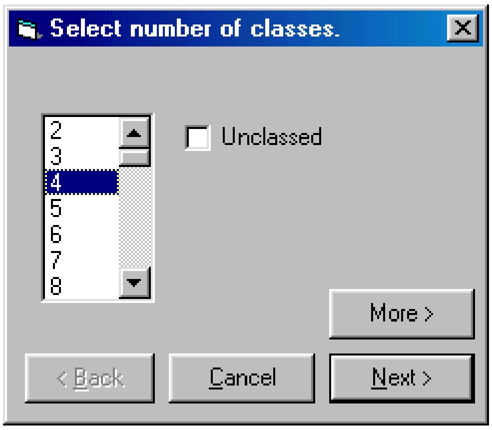
2. The choropleth mapping wizard

Among the variables listed in the *Select variable* window is the variable *ElderlyWelfare*, which is the number of places in institutions for the elderly per 1000 inhabitants above 67 years in Trøndelag in 1997. Select this variable by clicking on it.

GIB then investigates the values the selected variable holds and interprets them to be at relative ratio level. A class less choropleth map will then be generated as default.

Start the choropleth wizard (Ref: Figure 3.1.1 / 3.1.2)
- select *Choropleth map* from the *MapGallery* menu, or click on the *choropleth map* icon on the tool menu.

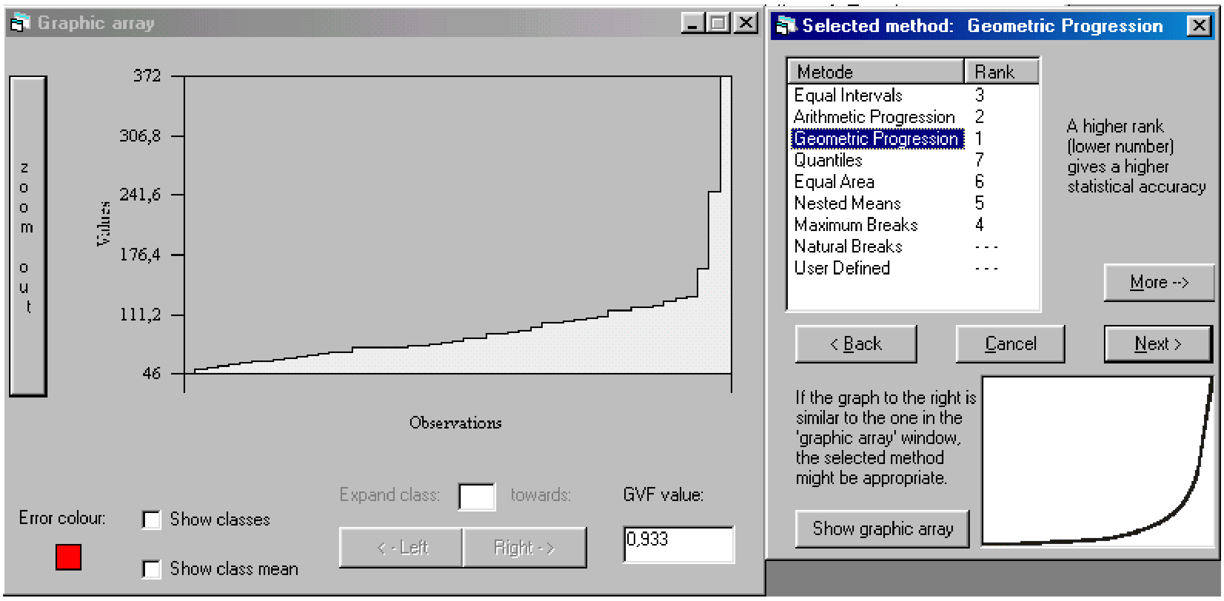
The map will than be regenerated as a base map and the *Select number of classes* dialog box will be loaded:



Select *four* classes
Click next

Ref: Figure 3.2

You will then be displayed the dialog box for selecting classification method.



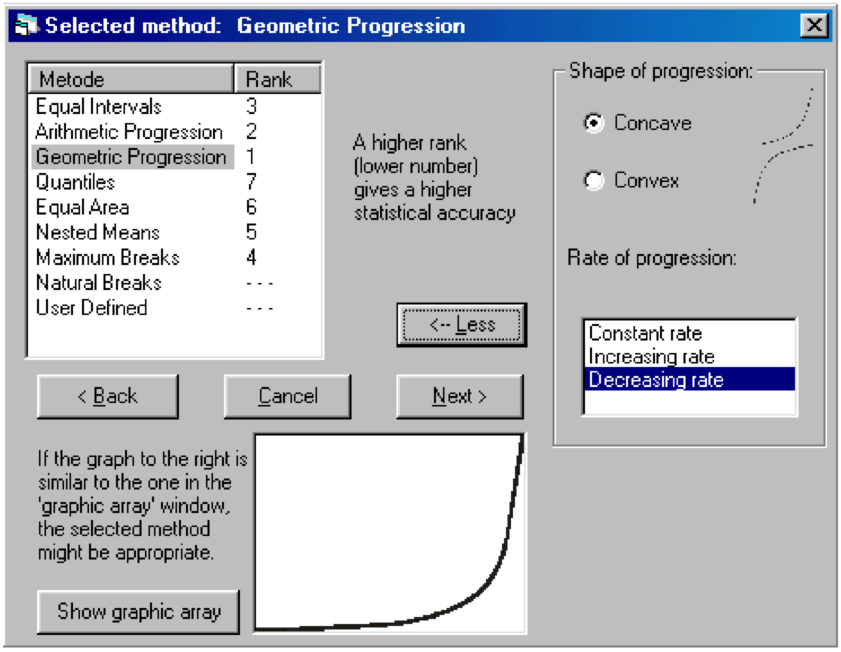
Metode	Rank
Equal Intervals	3
Arithmetic Progression	2
Geometric Progression	1
Quantiles	7
Equal Area	6
Nested Means	5
Maximum Breaks	4
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

If the graph to the right is similar to the one in the 'graphic array' window, the selected method might be appropriate.

(Ref: Figures 3.4, 3.5.1 and 3.52)

- The *Selected method* dialog box is the same as the one shown in Rød (VI, Figure 3a).
- Click on the *Show graphic array* button. A graphic array is then visible in a new window as shown above.
- Click on the *zoom to range* button.
- This variable is considerably skewed and consequently, the default method offered is a geometric progression. Note also that the method *standard deviation* is not included since this distribution deviates significantly from a normal distribution.
- Click on the *more button*.



Selected method: Geometric Progression

Metode	Rank
Equal Intervals	3
Arithmetic Progression	2
Geometric Progression	1
Quantiles	7
Equal Area	6
Nested Means	5
Maximum Breaks	4
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

Shape of progression:
● Concave
○ Convex

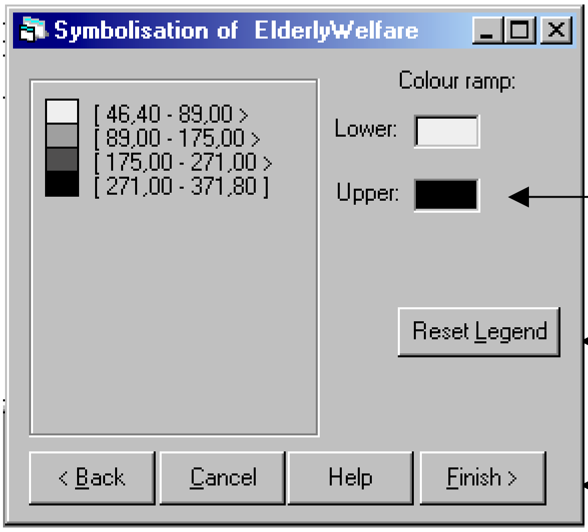
Rate of progression:
Constant rate
Increasing rate
Decreasing rate

If the graph to the right is similar to the one in the 'graphic array' window, the selected method might be appropriate.

Ref: Figure 3.6

GIB has already calculated the GVF values for the possible combinations of concave versus convex and whether the progression should have a constant, increasing or decreasing rate – the combination resulting in the highest GVF value is the default offered (see also Rød (VI: 11-12).

Click on *Next >* to load the dialog box for symbolisation:



Symbolisation of ElderlyWelfare

[46,40 - 89,00 >
[89,00 - 175,00 >
[175,00 - 271,00 >
[271,00 - 371,80]

Colour ramp:
Lower:
Upper:

Ref: Figure 3.8

Click on the *Upper* field to change the colour ramp.

Click on the *Reset Legend* to update the legend.

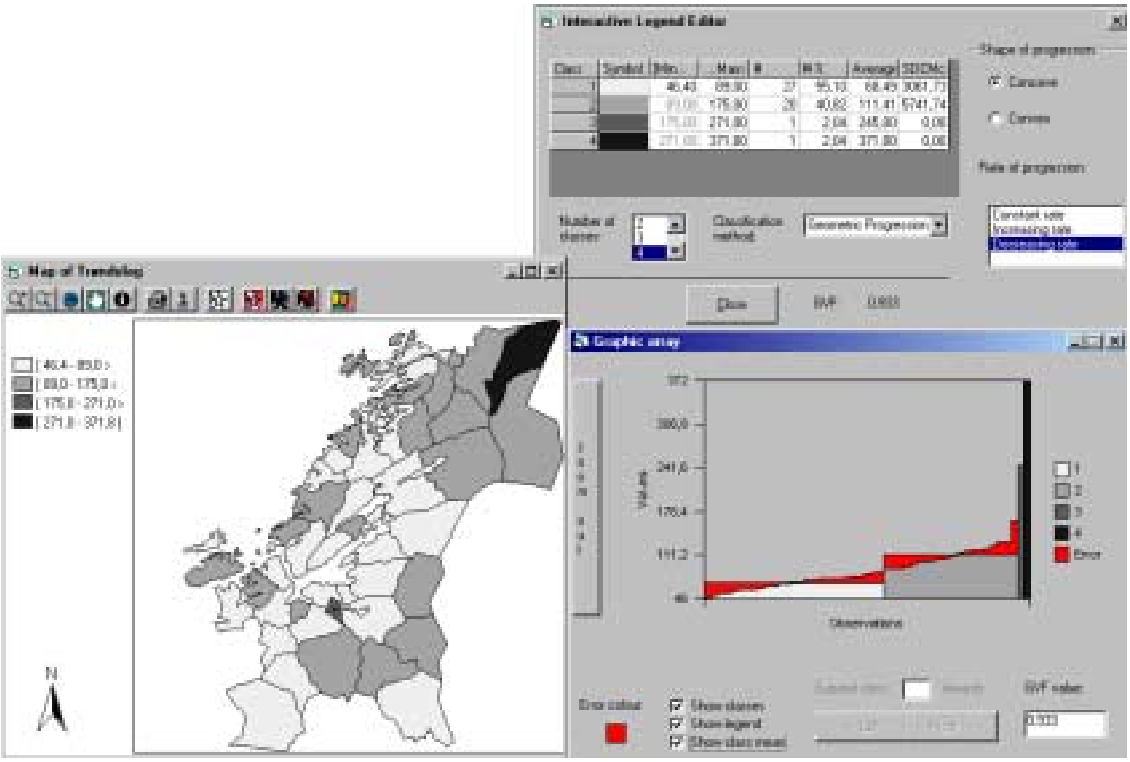
Click on Finish to update the map.

The map will then be symbolised in the colour ramp defined.

3. Linked windows

While the map still is displayed, launch the

- Interactive Legend Editor (from the Edit menu, select Edit Choropleth Map) and the
- Graphic Array (from the *View* menu, select *Graphic Array*)
- activate *show classes* and *show class means* and situate the dialog boxes on the screen along each sides. For instance something like this:



From the *Interactive Legend Editor*

- select *Quantiles* as classification method to produce a similar map, diagram and interactive legend editor as Figure 5 in Rød (VI:15).
- select *Equal Intervals* as classification method to produce a similar map, diagram and interactive legend editor as Figure 6 in Rød (VI:16).
- change the number of classes (from the *Interactive Legend Editor*) to see how the GVF values changes as shown in Figure 7 in Rød (VI: 18).

GIB

Geographic Information Processing

Manual

Jan Ketil Rød
Department of Geography
NTNU
21.06.2001

Content

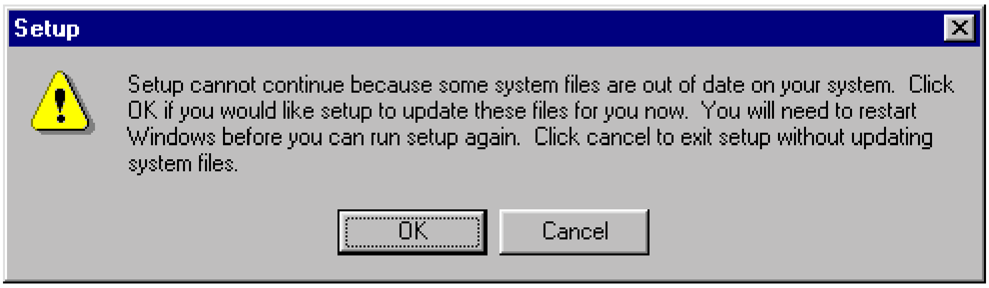
1. Installing GIB

1.1. Operative system requirements

Installation of GIB should be easy on Win95, Win98 and Win2000.

If Figure 1.2 is shown, the installation MUST be aborted.

1.2. Warning



Setup cannot continue because some system files are out of date on your system. Click OK if you would like setup to update these files for you now. You will need to restart Windows before you can run setup again. Click cancel to exit setup without updating system files.

Figure 1.2
If, when installing GIB, the above message is shown: Abort the installation (click cancel) – do not trust the setup to update your system files).

1.3. Start installation

Skriv inn navnet på programmet, mappen, dokumentet eller Internett-ressursen som skal åpnes.

Åpne: E:\GIB\Setup.exe

Figure 1.3.1
Start the installation by selecting *Start*, *Run*. From this dialog box, click on *Brows* in order to locate the setup.exe file at the GIB CD. Click OK and the installation will start.

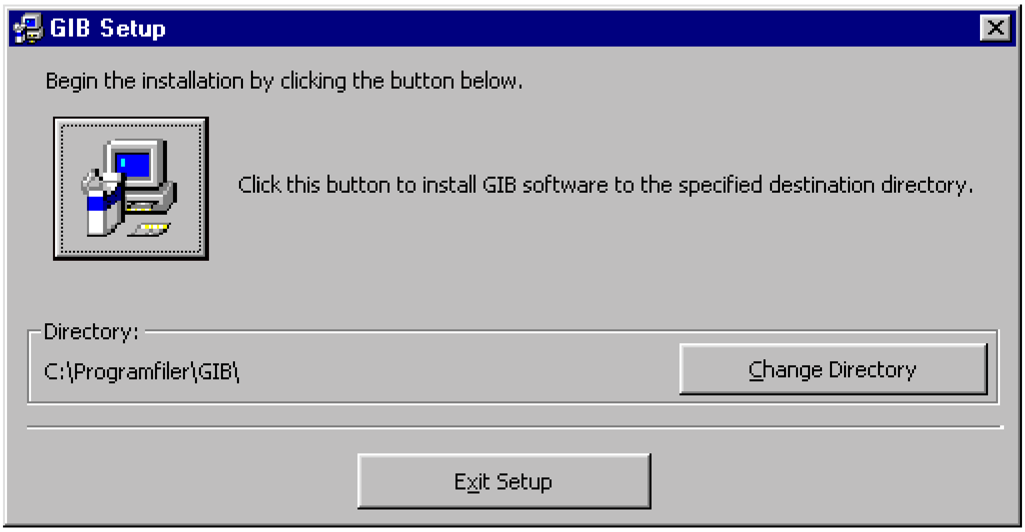
You can also locate the file by using *Explorer*. Double click on setup.exe and the installation starts.

Welcome to the GIB installation program.

Setup cannot install system files or update shared files if they are in use. Before proceeding, we recommend that you close any applications you may be running.

Figure 1.3.2
After launching the installation program, you are welcomed.
It is recommended to close all other running programs before running the setup.

1.4. Localisation



Begin the installation by clicking the button below.

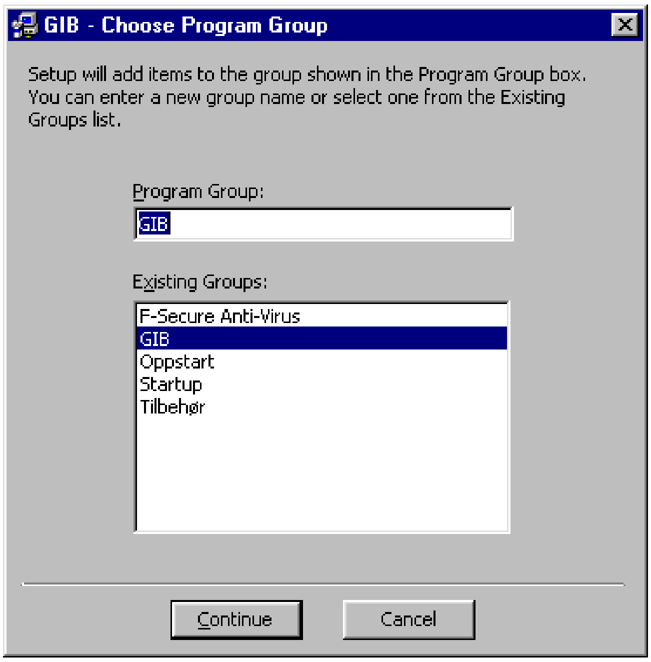
Click this button to install GIB software to the specified destination directory.

Directory:
C:\Programfiler\GIB\

Figure 1.4
If not expressed otherwise, GIB will be installed at the directory C:\Program files\GIB. Click on *Change directory* in order to install GIB at another directory. GIB is installed with the sub directory \Data. Along with GIB, are some data (statistics and maps), which might be used for tutorial purposes. New projects will also be stored under this directory. Click on the icon with the picture of a computer to start the installation.

The installation program will then test whether or not you have sufficient disk resources for the GIB program to be installed.

1.5. Select program group



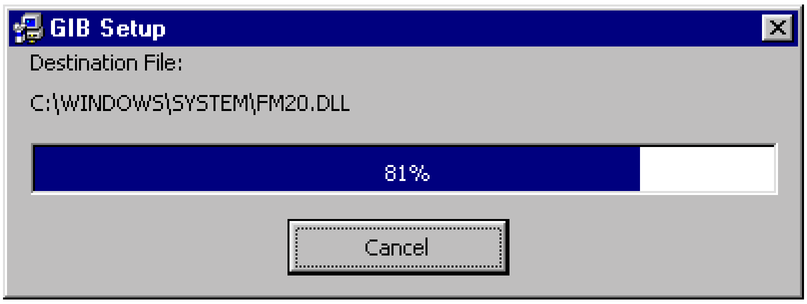
Setup will add items to the group shown in the Program Group box. You can enter a new group name or select one from the Existing Groups list.

Program Group: GIB

Existing Groups:
F-Secure Anti-Virus
GIB
Oppstart
Startup
Tilbehør

Figure 1.5
Click *Continue* if you agree on putting an icon in a program group called GIB (if not, change to whatever).

1.6. Installing



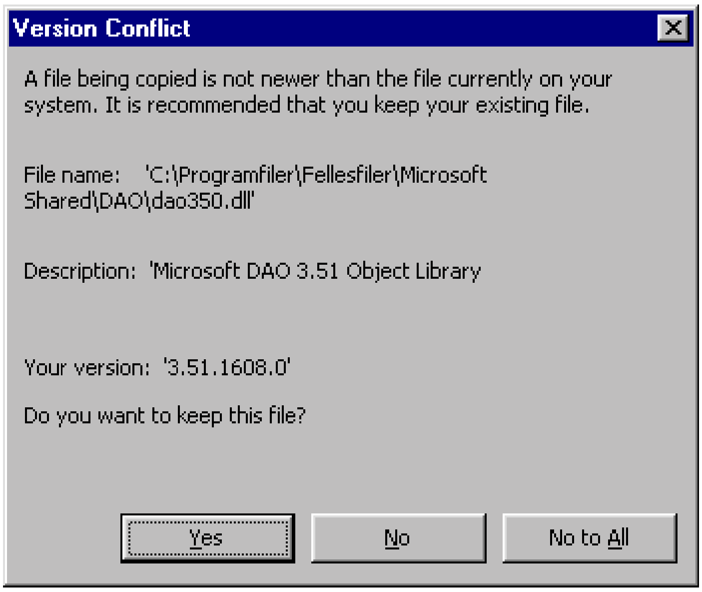
Destination File:

C:\WINDOWS\SYSTEM\FM20.DLL

81%

Figure 1.6
The installation of files will begin and you are prompted this dialog box showing which files being copied.

1.7. System files



A file being copied is not newer than the file currently on your system. It is recommended that you keep your existing file.

File name: 'C:\Programfiler\Fellesfiler\Microsoft Shared\DAO\dao350.dll'

Description: 'Microsoft DAO 3.51 Object Library

Your version: '3.51.1608.0'

Do you want to keep this file?

Figure 1.7
GIB uses several system files, which also other programs are using. Thus, you might already have stored on your computer several of the system files, which is needed to run GIB. If the system files, which you already have is older than those GIB is about to install, they will be replaced – if not, you are prompted this message. It is recommended that you do not replace new system files with old ones. Answer *Yes* to keep the existing files.

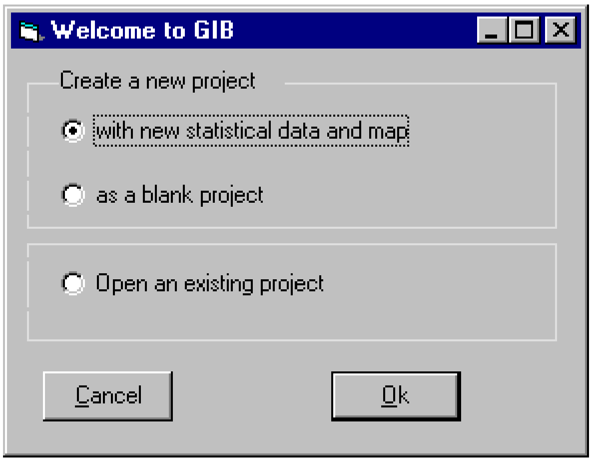
Thereafter, you will be prompting that GIB is updating your system and finally that the installation was successfully.

2. *Starting GIB*

Before being able to produce statistical maps with GIB, one needs an Excel file holding the statistical data and a boundary file in shape file format covering the units the statistical date are collected from. These data (statistics and map) must be loaded into GIB according to the procedure described here.

2.1. Welcome to GIB

Starting GIB you are first prompted a dialog box welcoming you to GIB (Figure 2.1).

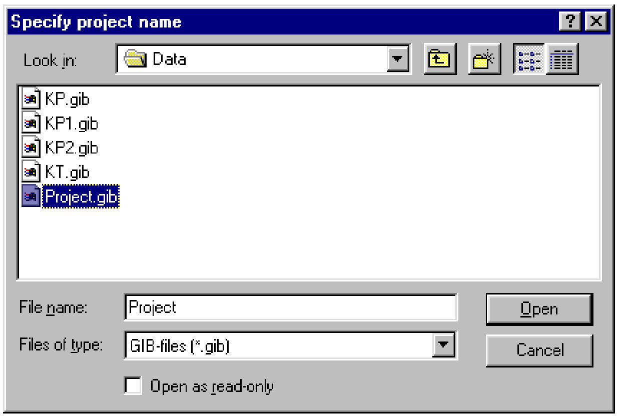


Create a new project
● with new statistical data and map
○ as a blank project
○ Open an existing project

Figure 2.1
The first dialog box after started GIB. You need to select between creating a new project (having two options: either subsequently loading statistical data and map or as a blank project (see section 2.3), or opening an existing project.

2.2. Open a project

Regardless which selection done in section 2.1, you need to open a project. A standard dialog box for specifying file name for opening these are shown (Figure 2.2).



Look in: Data

KP.gib
KP1.gib
KP2.gib
KT.gib
Project.gib

File name: Project
Files of type: GIB-files (*.gib)
☐ Open as read-only

Figure 2.2
The project is a file containing several parameters on how the statistical data shall be presented in a map. Select an existing project file or create a new one. If you are creating a new project and do not want to name it ”Project” you replace this name with your preference in the filename field.
For more details – see chapter 8 on storing.

2.3. Blank project

If you selected to start with a blank project (see 2.1.), the program loads empty with only some menus on the upper left corner (Figure 2.3.).

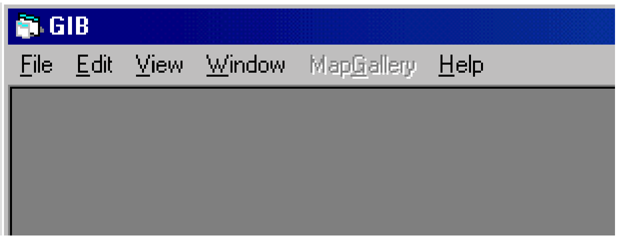
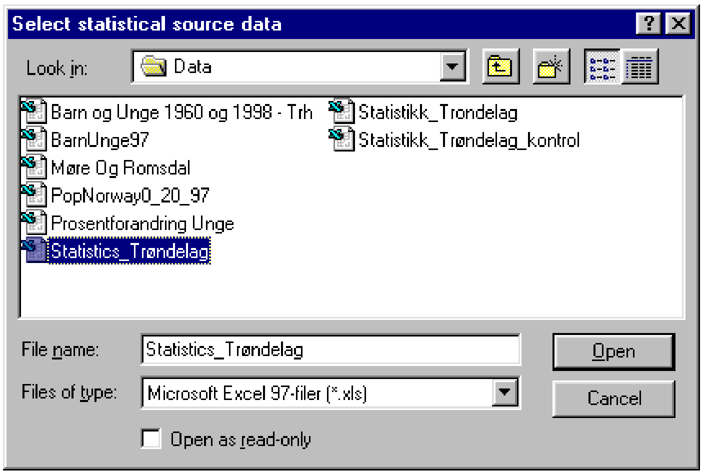


Figure 2.3
With a blank project – this menu line is shown in the upper left corner in the window opened. To open an Excel file containing statistics: select *File*, *Load Statistics*. To open a map: select *File*, *Load Base Map* (NB! You can only open a map if statistical data already is opened.)

2.4. Open statistical data

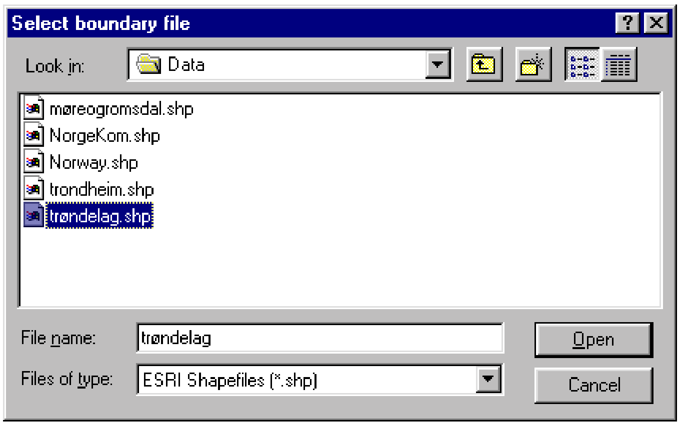
If you have selected to create a new project 'with new statistical data and map' (see Figure 2.1.) you will successive be prompted the dialog boxes shown in Figure 2.4 and 2.5. If you have selected to create a new project as a 'blank project' (see Figure 2.1.) you might open statistical data by selecting *File*, *Load Statistics*. The following dialog box will be shown:



File name: Statistics_Trøndelag
Files of type: Microsoft Excel 97-filer (*.xls)
☐ Open as read-only

Figure 2.4
Open statistical data.
GIB reads Excel files. This file must be adapted as described in appendix 1.

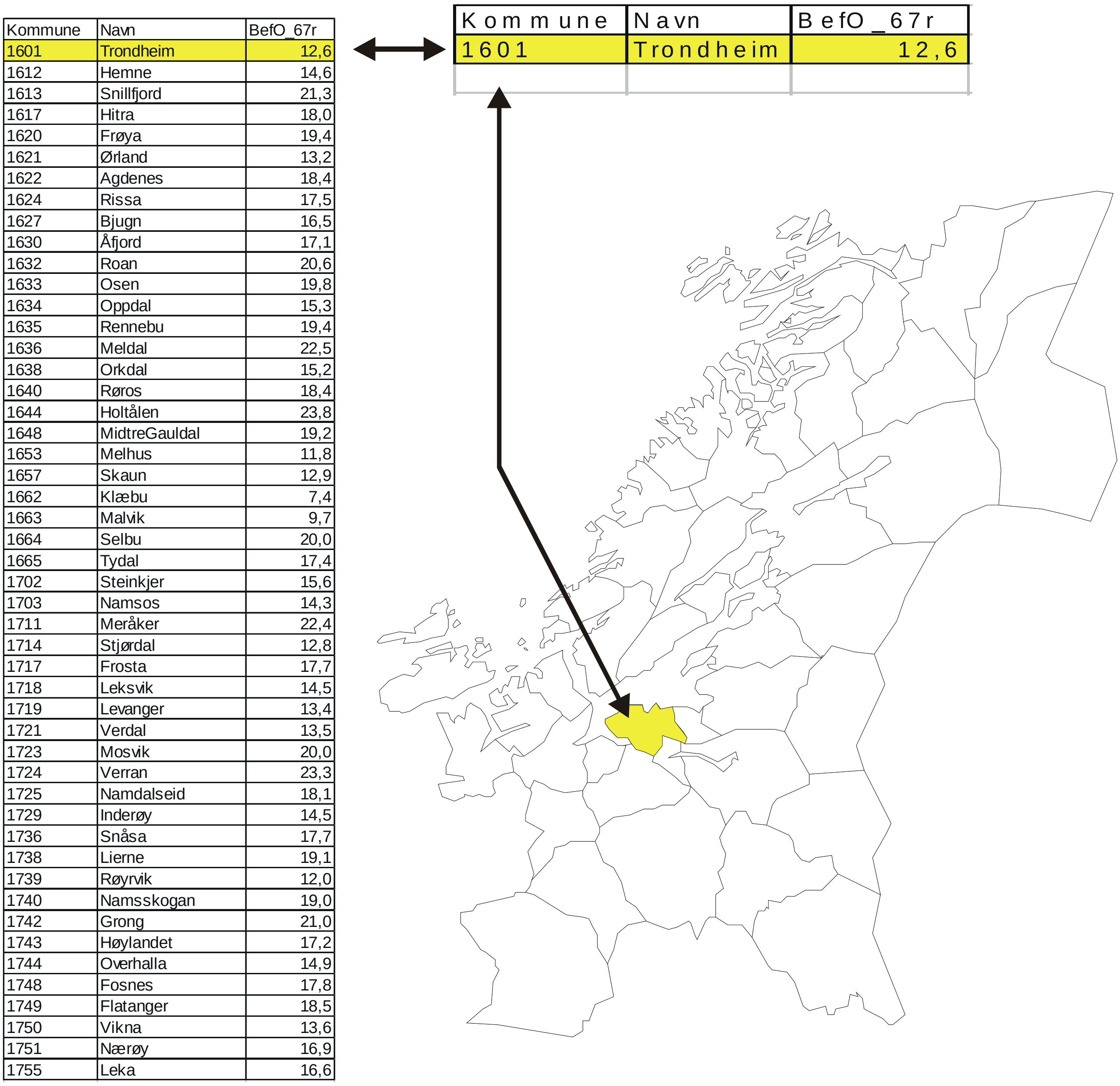
2.5. Open a map



File name: trøndelag
Files of type: ESRI Shapefiles (*.shp)

Figure 2.5
Open a map.
GIB reads map files on the 'Shape' format (see appendix 2).
If you opened a blank project (see 2.1.) this dialog box will be loaded by selecting *File*, *Load Base Map*.

2.6. Joining the Excel file with the map file

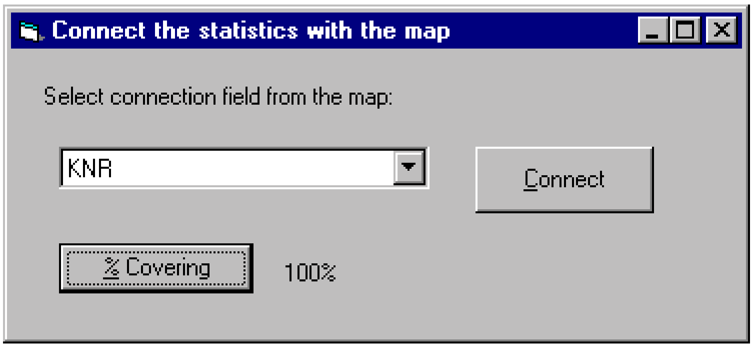


Kommune	Navn	BefO_67r
1601	Trondheim	12,6
1612	Hemne	14,6
1613	Snillfjord	21,3
1617	Hitra	18,0
1620	Frøya	19,4
1621	Ørland	13,2
1622	Agdenes	18,4
1624	Rissa	17,5
1627	Bjugn	16,5
1630	Åfjord	17,1
1632	Roan	20,6
1633	Osen	19,8
1634	Oppdal	15,3
1635	Rennebu	19,4
1636	Meldal	22,5
1638	Orkdal	15,2
1640	Røros	18,4
1644	Holtålen	23,8
1648	MidtreGauldal	19,2
1653	Melhus	11,8
1657	Skaun	12,9
1662	Klæbu	7,4
1663	Malvik	9,7
1664	Selbu	20,0
1665	Tydal	17,4
1702	Steinkjer	15,6
1703	Namsos	14,3
1711	Meråker	22,4
1714	Stjørdal	12,8
1717	Frosta	17,7
1718	Leksvik	14,5
1719	Levanger	13,4
1721	Verdal	13,5
1723	Mosvik	20,0
1724	Verran	23,3
1725	Namdalseid	18,1
1729	Inderøy	14,5
1736	Snåsa	17,7
1738	Lierne	19,1
1739	Røyrvik	12,0
1740	Namsskogan	19,0
1742	Grong	21,0
1743	Høylandet	17,2
1744	Overhalla	14,9
1748	Fosnes	17,8
1749	Flatanger	18,5
1750	Vikna	13,6
1751	Nærøy	16,9
1755	Leka	16,6

Kommune	Navn	BefO_67r
1601	Trondheim	12,6

Figure 2.6.1.
Joining a table with a map in GIB require that each geographical unit is represented by a row in the data table and that the first row is a common identifier.

In order to visualise statistical data stored in an Excel spreadsheet cartographically, the Excel file must be joined with a map, which is possible by a common identifier. In the Excel spreadsheet, this is the first column containing, for instance, the identification numbers for the municipalities (i.e. for Trondheim: 1601). If the column heading for this connection field is identical for both the Excel file and the map, joining them will be done automatically. If not, the column in the map file, which should be used to match the first column in the Excel spreadsheet, must be indicated from the dialog box shown in Figure 2.6.2.



Connect the statistics with the map

Select connection field from the map: KNR

% Covering 100%

Figure 2.6.2
Select connection field (here: KNR) for manual connecting the Excel spreadsheet to the map. To sure you have selected the right connection field, click on *%covering* to check the percentage covered. Click on *Connect* to join.

2.7. Map window

When Excel spreadsheet file and map is opened (and joined), the map will be shown.

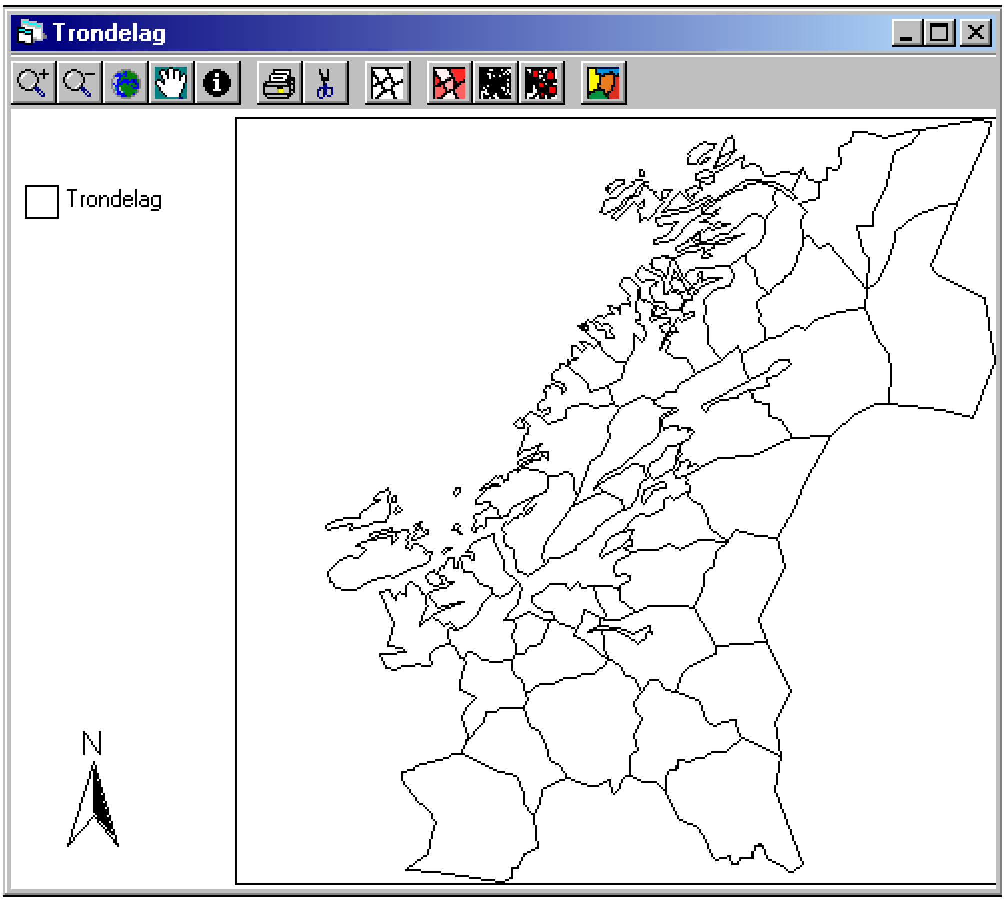
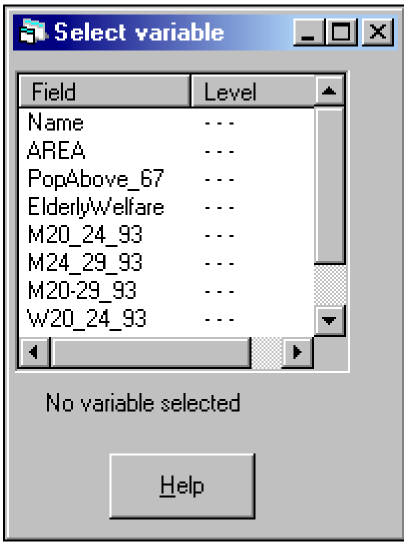


Figure 2.7
The map window shows the map, legend (left part), a north arrow and some menu buttons. The base map shown here covers the two counties South– and North Trøndelag. The base map might be symbolised according to one of the variables from the spreadsheet opened. Available types of maps are choropleth map, dot density map and point proportional symbol map and nominal map (chorochromatic map).

2.8. Select variable



Field	Level
Name	---
AREA	---
PopAbove_67	---
ElderlyWelfare	---
M20_24_93	---
M24_29_93	---
M20-29_93	---
W20_24_93	---

No variable selected

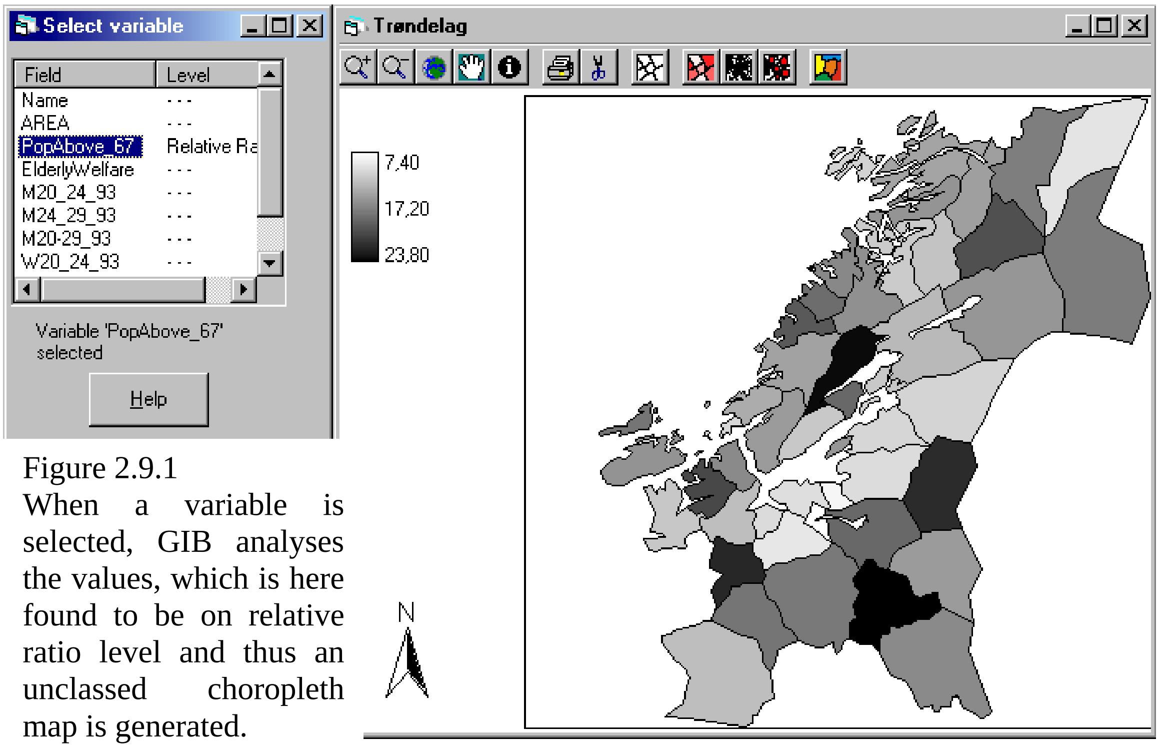
Figure 2.8
After having indicated the spreadsheet file to open (Figure 2.4) the variable names (the column headings), which is present in the spreadsheet will be listed in a window called ‘Select variable’. As long as no variable is selected, the message “No variable selected” is shown and the field “level”, which indicate the variables level of measurement, will have the value “---”.

Click on a variable name to select it.

2.9. Determine measurement level

Immediately after a variable is selected, a default map is generated depending on the level of measurement. The assignment of map type to measurement level follows this scheme:

nominal values	Ą	nominal map (chorochromatic map)
ordinal values	Ą	unclassed choropleth map
interval values	Ą	unclassed choropleth map
relative ratio	Ą	unclassed choropleth map
absolute ratio	Ą	point proportional symbol map

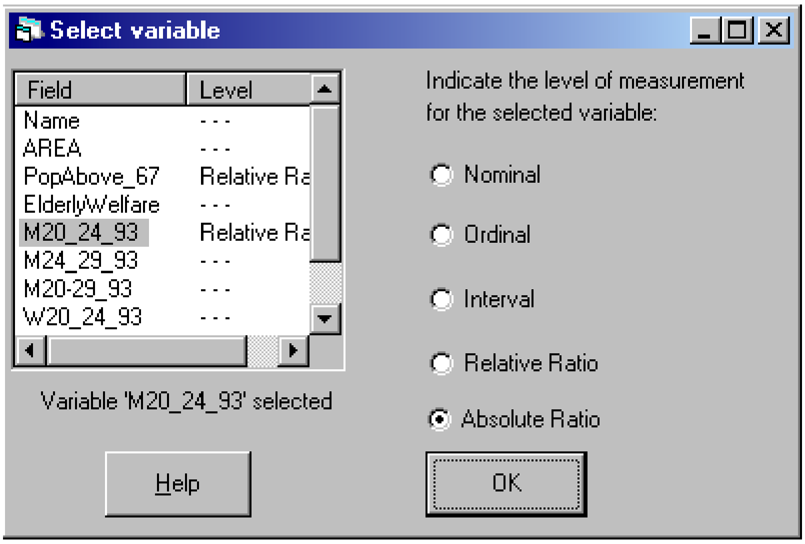


Field	Level
Name	---
AREA	---
PopAbove_67	Relative Ra
ElderlyWelfare	---
M20_24_93	---
M24_29_93	---
M20-29_93	---
W20_24_93	---

Variable 'PopAbove_67' selected

Figure 2.9.1
When a variable is selected, GIB analyses the values, which is here found to be on relative ratio level and thus an unclassed choropleth map is generated.

If you, for some reason need to re-indicate the level of measurement, select *File* and *Edit Measurement Level*.



Field	Level
Name	- - -
AREA	- - -
PopAbove_67	Relative Ra
ElderlyWelfare	- - -
M20_24_93	Relative Ra
M24_29_93	- - -
M20-29_93	- - -
W20_24_93	- - -

Variable 'M20_24_93' selected

Indicate the level of measurement for the selected variable:

○ Nominal
○ Ordinal
○ Interval
○ Relative Ratio
● Absolute Ratio

Figure 2.9.2
The variable *M20_24_93* (Total number of men aged between 20 and 24 in the Trøndelag region in 1993) is set to relative ratio level, but should be on absolute ratio level. Click on the option *Absolute Ratio* to change.

3. Choropleth map

Three decisions must be made when producing a choropleth map:
- Selecting the number of classes
- Selecting method for class interval determination
- Selecting symbolisation

In GIB you are lead through these three choices by a wizard. When all three choices are made a choropleth map, which you later may change interactively (see chapter four), is generated.

3.1. Starting the choropleth wizard

To make a choropleth map in GIB:

Figure 3.1.1
Select Choropleth map from Map Gallery.

or

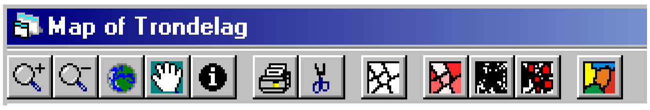


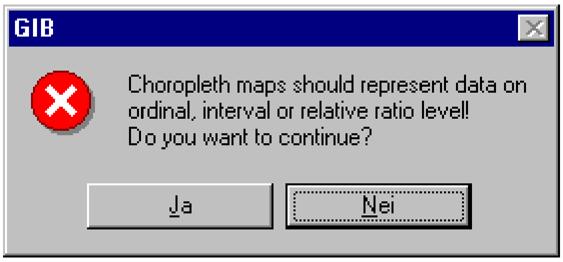
Figure 3.1.2
Click on the choropleth map icon.

One cannot make a choropleth map without having selected a variable, in which case you are prompted the message shown in Figure 3.1.3.

You need to choose variable before making the choropleth map!

Figure 3.1.3
To make a choropleth map, a variable must be selected.

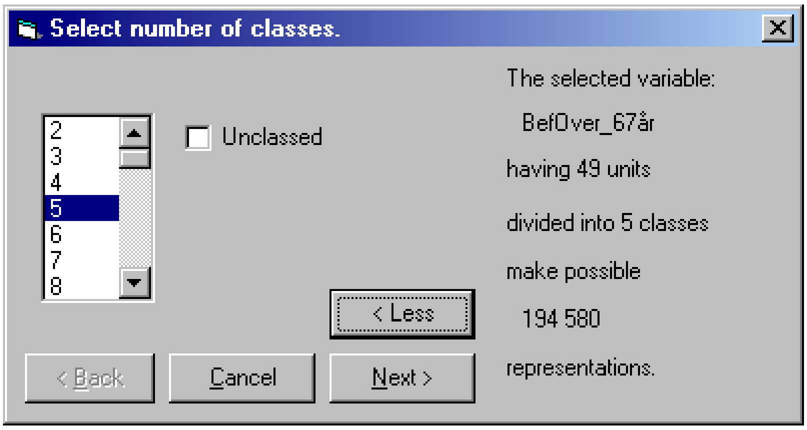
To make a choropleth map, the variable ought to be on relative ratio level. If you have selected a variable, which has not this level indicated, you will be prompted the message shown in Figure 3.1.4.



Choropleth maps should represent data on ordinal, interval or relative ratio level!
Do you want to continue?

Figure 3.1.4
Warning against using a representational form not suited for the indicated level of measurement.

3.2. Selecting the number of classes



2
3
4
5
6
7
8

☐ Unclassed

The selected variable:

BefOver_67år

having 49 units

divided into 5 classes

make possible

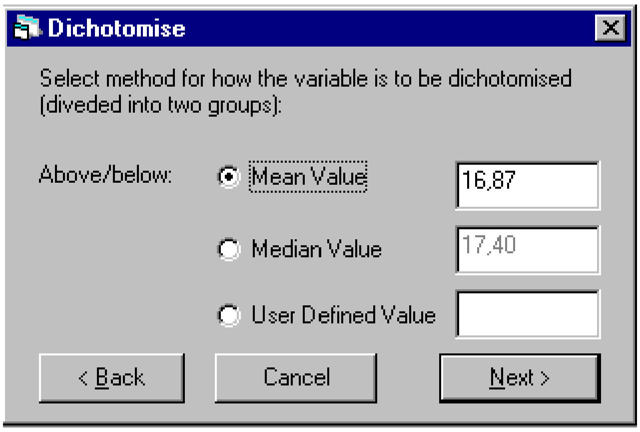
194 580

representations.

Figure 3.2
By clicking on ”More >” you will have more information on how many different choropleth maps it is possible to make from the selected variable divided into the selected number of classes.

The number of classes can be between two and a number equal the number of unique values. By selecting two, a dichotomised map will be made (see section 3.3.), if not, an ordinary choropleth map will be produced (see section 3.4 and further). If the check box titled *Unclassed* is marked, an unclassed choropleth map will be generated.

3.3. Number of classes equals two



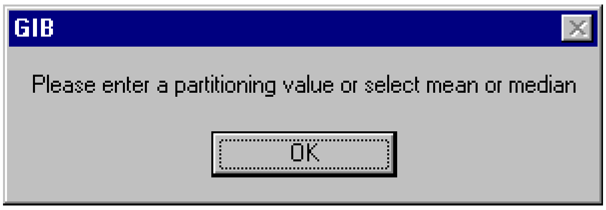
Select method for how the variable is to be dichotomised (diveded into two groups):

Above/below: ● Mean Value 16,87

○ Median Value 17,40

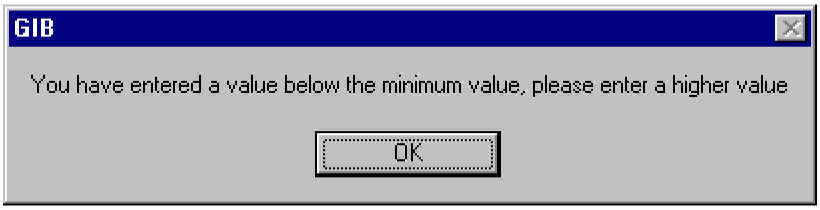
○ User Defined Value

Figure 3.3.1
If you have selected ‘two’ on the number of classes, the data set will be divided into two groups, respectively above or below the arithmetic mean, the median or another user defined value.



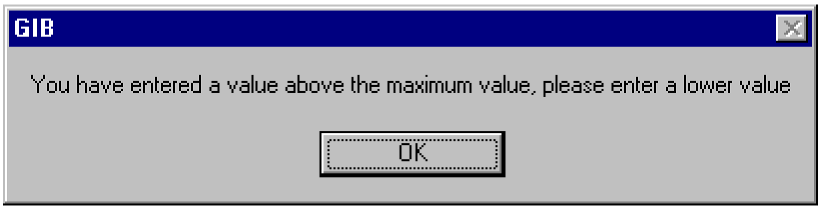
Please enter a partitioning value or select mean or median

Figure 3.3.2
If you have selected to enter a user defined value you must enter this value in the open field to the right for “User Defined Value”. If you click on “Next >” without having entered this value, you are prompted this message.



You have entered a value below the minimum value, please enter a higher value

Figure 3.3.3
The value you have entered is too low.



You have entered a value above the maximum value, please enter a lower value

Figure 3.3.4
The value you have entered is too high.

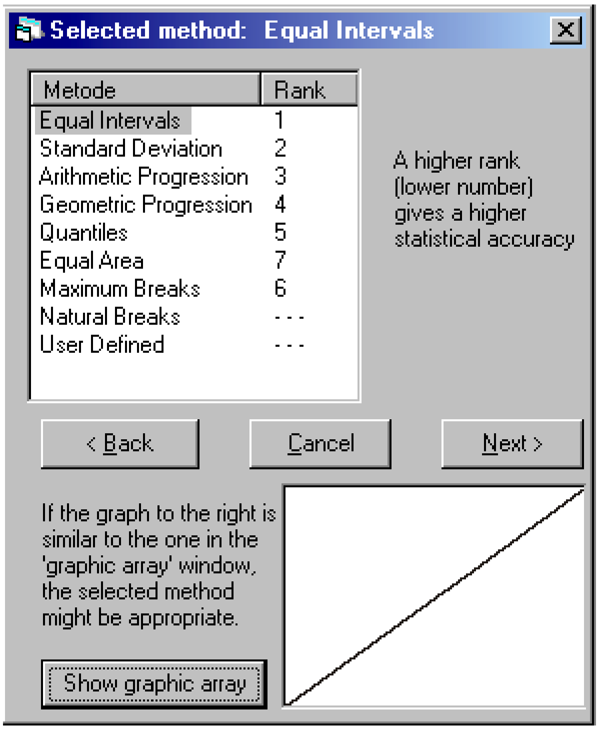
3.4. Selecting classification method

There are several methods for classifying a variable. Usually, there is in mapping software packages offered only a limited number of methods. GIB offers in total ten methods (see Cauvin et al 1987, Dent 1996, Robinson et al 1984, Slocum 1999 for an outline of the various methods). The high numbers of classification methods available makes the choice more complicated and it is thus implemented resources, which determine the most appropriate method for the selected variable grouped in the selected number of classes. The available resources for decision support depend upon type of classification. There are three sorts of classification:

Serial	Equal Intervals	
	Standard Deviation	(see section 3.6)
	Arithmetic Progression	(see section 3.7)
	Geometric Progression	(see section 3.7)
Idiographic	Quantiles	
	Equal Areas	
	Nested Means	
	Maximum Breaks	
	Natural Breaks	(see section 3.8)
External	User Defined	(see section 3.9)

For the serial methods, an ideal distribution is shown in a small window (see Figure 3.4). Click on *Show graphic array* to show the selected variable's distribution. If the selected variable's distribution resembles the ideal one, the actual method is appropriate

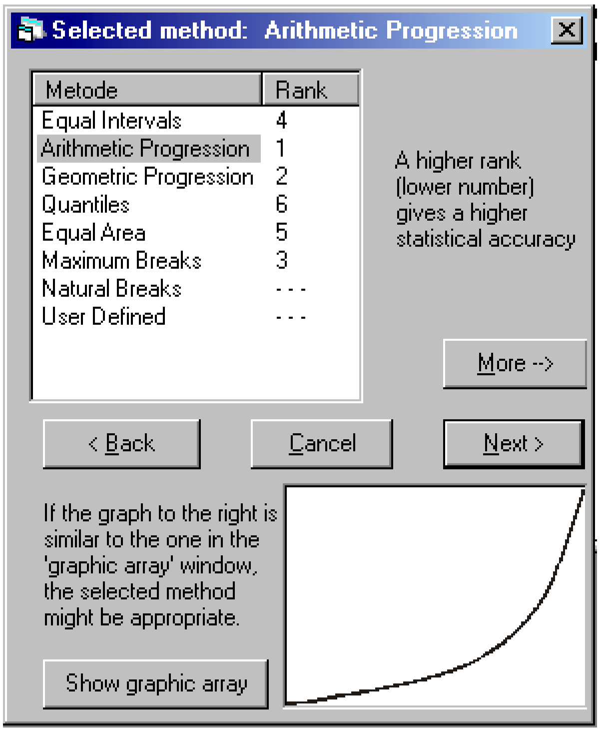
For the serial and idiographic methods, an index called GVF is calculated. The GVF value indicates the methods suitability for the selected variable grouped into the selected number of classes. Each of the methods is ranked according to its score on this index. The higher rank (the less number in the rank column), the more suitable is the method. The method with rank 1 is the default offered.



Metode	Rank
Equal Intervals	1
Standard Deviation	2
Arithmetic Progression	3
Geometric Progression	4
Quantiles	5
Equal Area	7
Maximum Breaks	6
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

If the graph to the right is similar to the one in the 'graphic array' window, the selected method might be appropriate.



Metode	Rank
Equal Intervals	4
Arithmetic Progression	1
Geometric Progression	2
Quantiles	6
Equal Area	5
Maximum Breaks	3
Natural Breaks	- - -
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

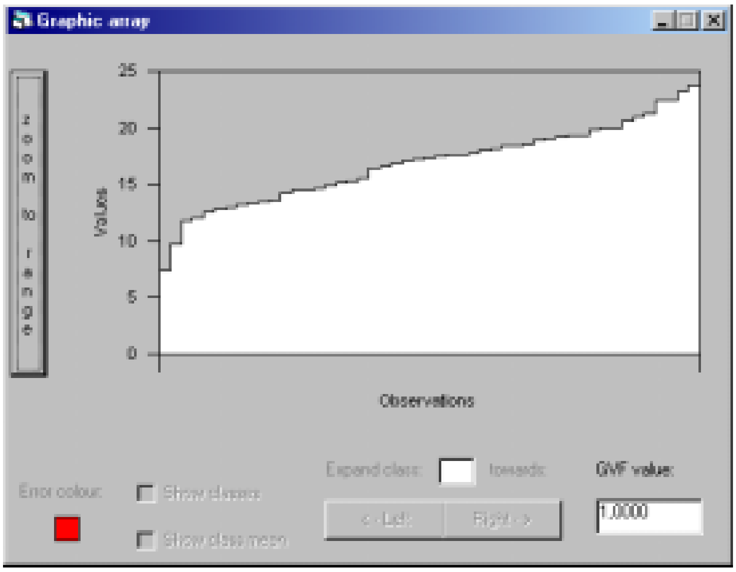
If the graph to the right is similar to the one in the 'graphic array' window, the selected method might be appropriate.

Figure 3.4
The number of available methods will vary. If the variable's distribution deviates significant from a normal distribution, the 'standard deviation' method will not be included among the options (See Rød, VI). If number of classes do not equal 4, 8, 16, or 2^n then the 'nested means' method is excluded. For the serial methods, a visual comparison between the ideal and actual distribution can be done. The actual distribution (see Figure 3.5) will be shown by clicking on *Show graphic array*. (For the methods arithmetic and geometric progression, there are several parameters which may be altered (see Figure 3.6), these are shown by clicking the button *More* (the button is only visible if one of these methods is selected)).

3.5. Graphic array

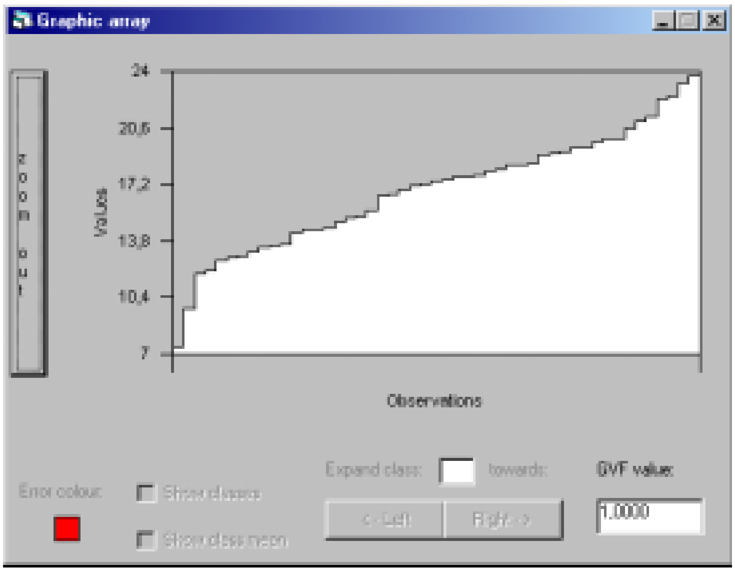
The selected variable's actual distribution is shown by loading the graphic array. The heights of the bars represent the observation values, which are sorted in increasing order

Figure 3.5.1
When a variable is selected (and if this variable do not hold values on nominal level) a graphic array of sorted observation values can be shown by selecting *View, Graphic Array*.



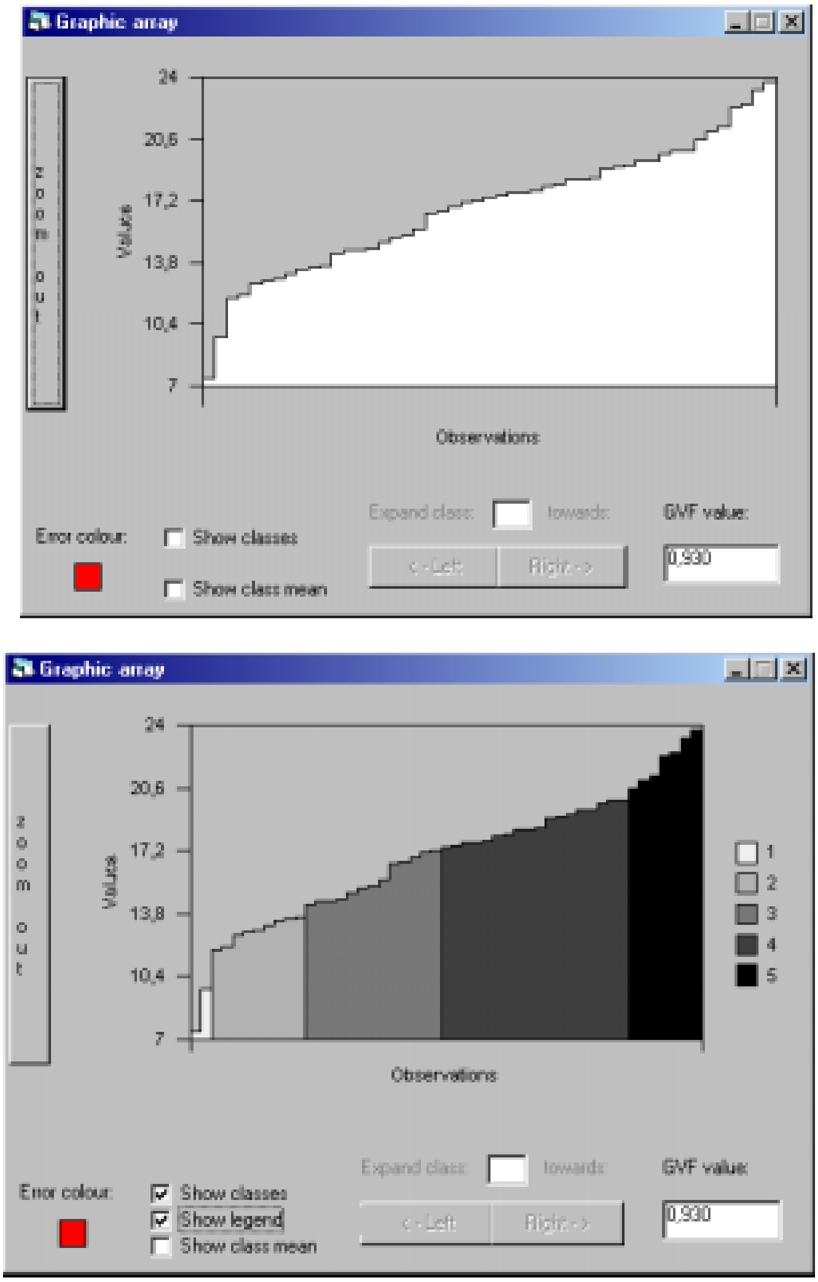
GVF value: 1,0000

By clicking on *zoom to range* the scale of the second axis is adjusted to the variable's range (the text on the button changes to *zoom out*).



GVF value: 1,0000

If no classification is performed, the options on the lower part of the dialog box (e.g. *show classes*) are grey shaded and the GVF index equals one.



GVF value: 0,930

☐ Show classes
☐ Show class mean

GVF value: 0,930

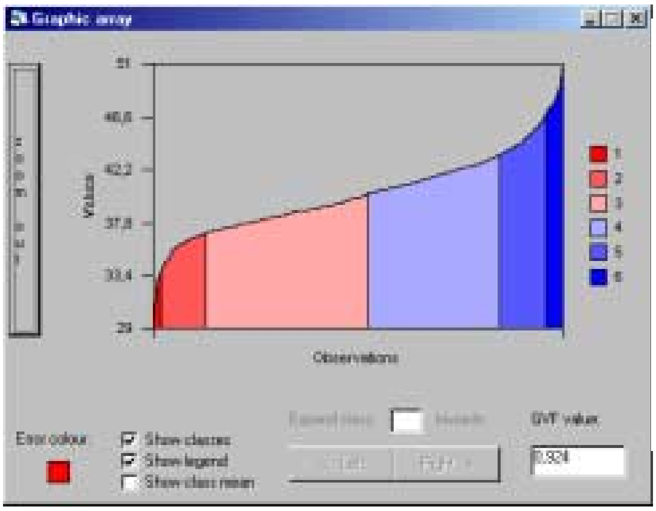
☑ Show classes
☑ Show legend
☐ Show class mean

Figure 3.5.2
The graphic array can also be loaded when one is about to select method for data classification by clicking on the *Show graphic array* button (see Figure 3.4). The graphic array displays the calculated GVF index based on the number of classes and the classification method selected.

The *show classes* option is enabled and by activating it, the bars are coloured according to the class they are grouped into. The legend can be turn on/off as well as class means and the observations deviations from class means. For more details on these and related functionalities, see chapter 4.

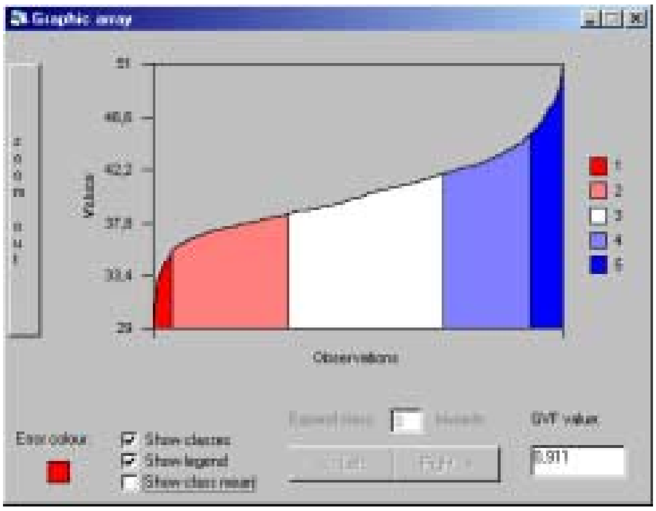
3.6. Standard deviation

The option on using standard deviation as classification method is available only if the selected variable does not deviate too much from a normal distribution (see Rød, VI). If the standard deviation option is available and selected, the class width will be one standard deviation except for the extreme lower and upper classes. If you have selected an even number of classes, classes below the arithmetic mean will be symbolised with a shade of red with an increased intensity proportional with an increased distance from the mean and classes above the arithmetic will by symbolised similarly, but with shades of blue. If you have selected an odd number of classes, the central class (± ½ standard deviation from the mean) will be coloured white while the lower and upper classes will be coloured with shades of red and blue respectively as with even number of classes. Below are two graphic arrays for the variable *WM0_20_97r* (percentage of the population aged 0 to 20 years in Norway in 1997) divided into six and five classes respectively according to the standard deviation method.



☑ Show classes
☑ Show legend
☐ Show class mean
GVF value 0,924

Figure 3.6.1
An even number of classes are selected (six classes); three classes below and three classes above the arithmetic mean (40,23).
The distance from the mean to the extreme lower class and the extreme upper class are two standard deviations.

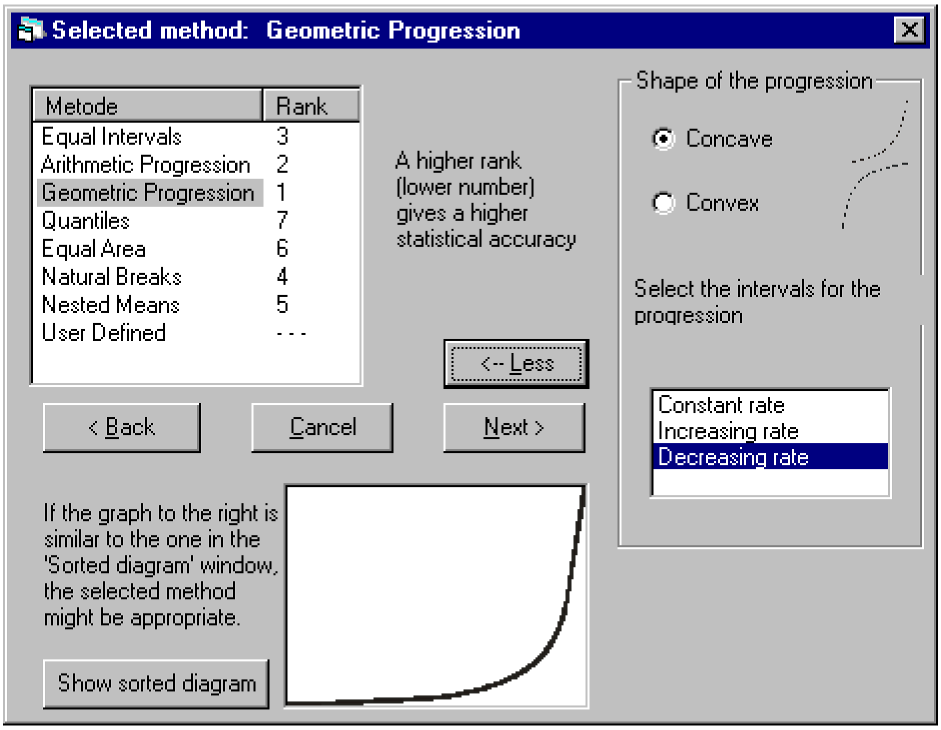


☑ Show classes
☑ Show legend
☐ Show class mean
GVF value 0,911

Figure 3.6.2
An odd number of classes are selected; one central class, two below and two above the central class.
The distance from the mean to the extreme lower class and the extreme upper class are one and a half standard deviations.

3.7. Arithmetic and geometric progression

Arithmetic and geometric progressions may take different forms. They might be concave or convex and the progression might increase with a constant, increasing or decreasing rate. The complexity concave versus convex entail is attempted to be simplified by visual tools and by reasonable default options: The one of the six alternatives resulting in the best GVF value is marked as default.



Metode	Rank
Equal Intervals	3
Arithmetic Progression	2
Geometric Progression	1
Quantiles	7
Equal Area	6
Natural Breaks	4
Nested Means	5
User Defined	- - -

A higher rank (lower number) gives a higher statistical accuracy

Shape of the progression

● Concave

○ Convex

Select the intervals for the progression

Constant rate
Increasing rate
Decreasing rate

If the graph to the right is similar to the one in the 'Sorted diagram' window, the selected method might be appropriate.

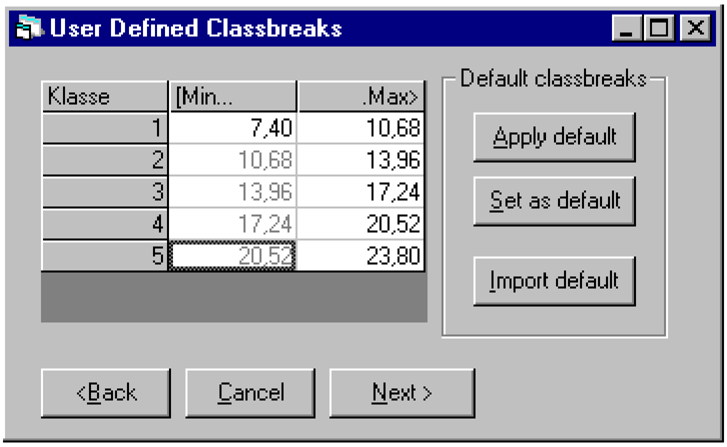
Figure 3.7
The methods arithmetic and geometric progression have additional options. The form the progression can be either concave or convex and the intervals for the progressions might increase with a constant, increasing or decreasing rate. In total, this gives six possibilities. The combination, which results in the most adequate method (best GVF value), is set as default.

3.8. Natural breaks

'Natural breaks' is a graphic method where the class breaks are determined by using a histogram interactively. The histogram is loaded automatically when selecting 'natural breaks' as classification method. See chapter 7.2 for details on functionality regarding the histogram.

3.9. User defined class breaks

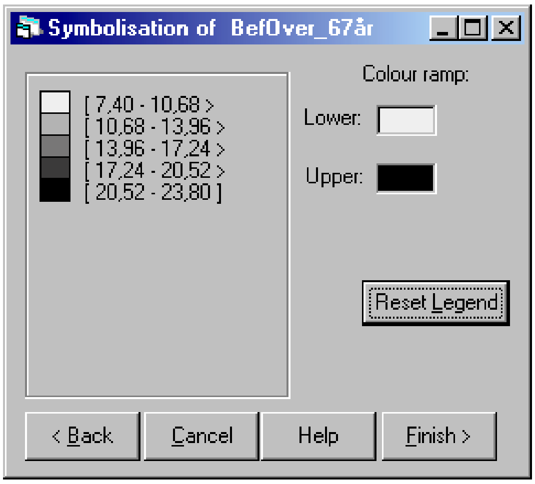
If a classification scheme is to be used for several maps, which might be relevant when making maps for time series, class breaks may be set default by clicking on the *Set as default* button in the *User Defined Classbreak* dialog box (see Figure 3.9). The next time the classification scheme is used, the default classification scheme may be loaded by clicking on the *Apply default* button. Use the *Import default* button if the classification scheme is saved in another project.



Klasse	[Min...	.Max>
1	7,40	10,68
2	10,68	13,96
3	13,96	17,24
4	17,24	20,52
5	20,52	23,80

Figure 3.9
User defined class breaks can be set by clicking on the cells whose values you want to alter. Only the first class' minimum value, but all maximum values can be edited. If a maximum value is changed the next class' minimum value updates automatically.

3.10. Selecting symbolisation



[7,40 - 10,68 >
[10,68 - 13,96 >
[13,96 - 17,24 >
[17,24 - 20,52 >
[20,52 - 23,80]

Figure 3.10.1
Symbolising a choropleth map is by default a ramp from light (near white) to dark (black). To change the lower or upper colour in the ramp, click on *Lower* or *Upper* respectively (see Figure 3.10.2. for further instructions). To update the scheme, click on *Reset Legend*.

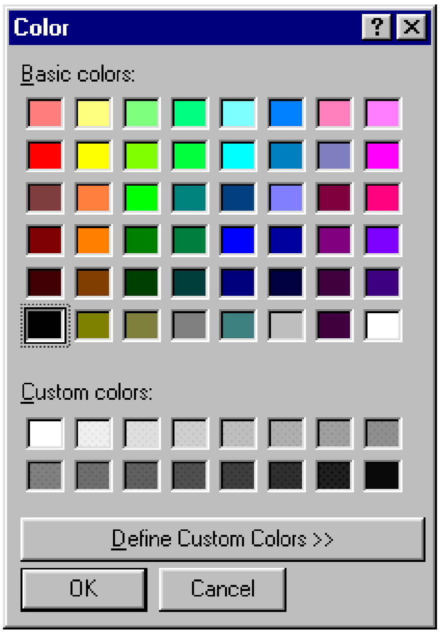


Figure 3.10.2
Click on one of the basic colour or define a custom colour to update the lower or upper colour in the ramp.

You will have several more options in defining your own colours if you click on the "*Define Custom Colors >>*".

Click on *OK* to confirm.

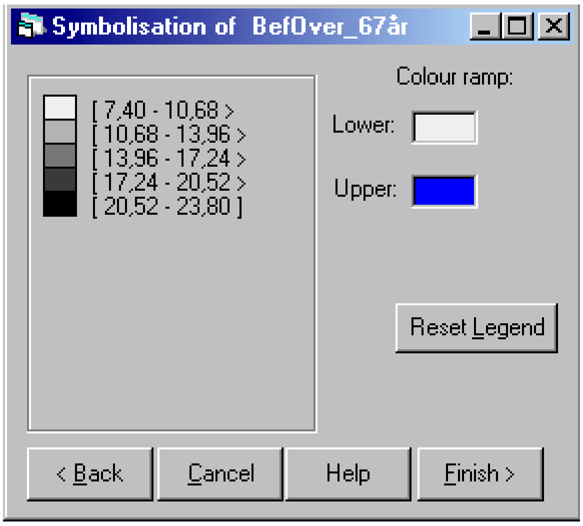


Figure 3.10.3
If you in Figure 3.10.2 have selected a blue colour for the upper colour, this will be the result.

To define a ramp from light to dark blue, click on *Reset Legend* (the result is shown in Figure 3.10.4).

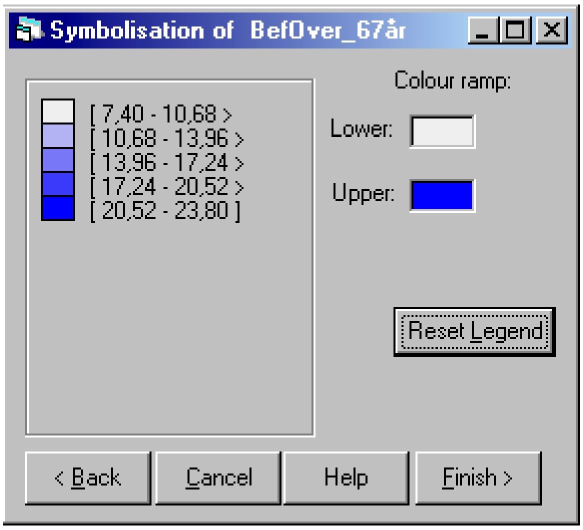


Figure 3.10.4
The colour ramp is updated in a light to dark colour range in blue.

3.11. Bipolar symbolisation

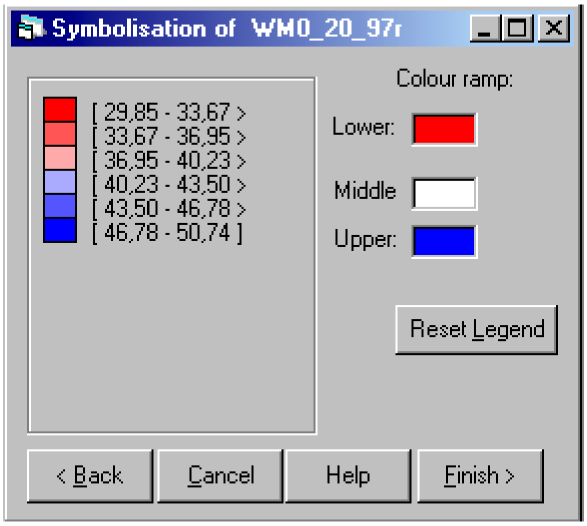


Figure 3.11
When the method 'Standard Deviation' is used the symbolisation scheme will be bipolar, that is a double colour range respectively above and below a centre (which is the arithmetic mean by the standard deviation method).

Changes (e.g. swapping red by blue) can be done as explained above.

3.12. Symbolisation for class less choropleth maps

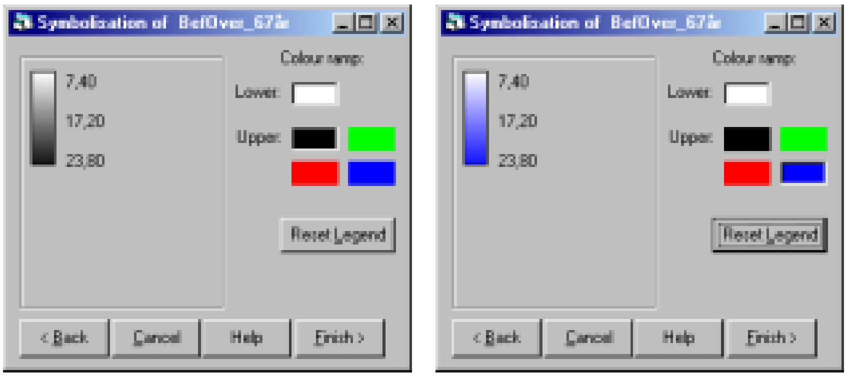


Figure 3.12
Symbolisation for a class less choropleth map is more restricted regarding choice of colours.

3.13. The map updates

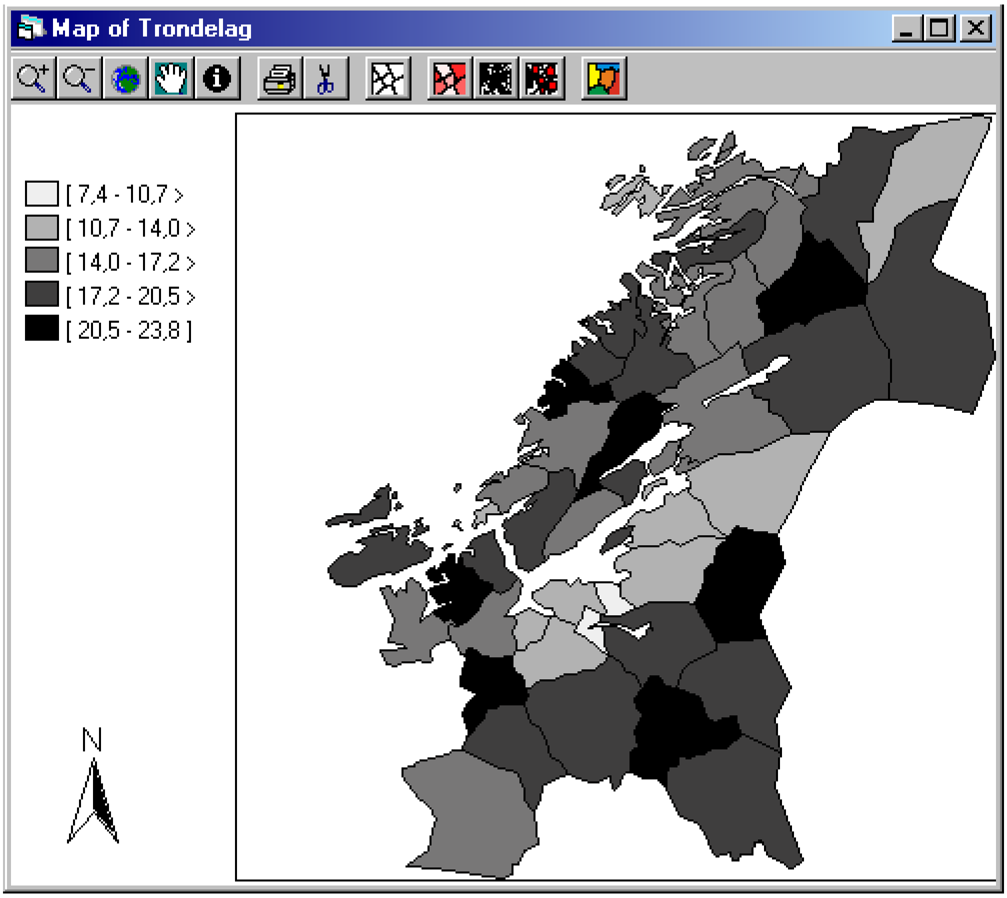


Figure 3.13
Choropleth map showing the percentage of elderly people (above 67 years) in the two counties, South- and North Trøndelag, in 1997. The variable is divided into five classes according to the equal interval method.

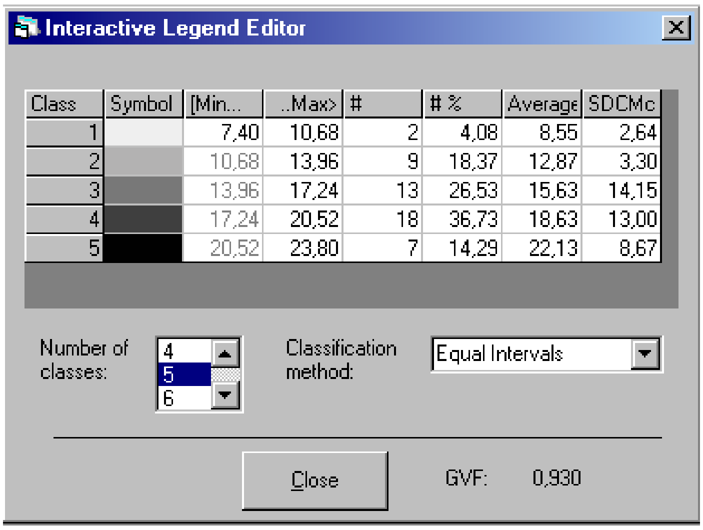
4. *Interactive editing and evaluating the choropleth map*

The map is not made ones for all in GIB. You may evaluate it or change it by using the *Interactive Legend Editor*.

4.1. Interactive legend editor

Figure 4.1.1
To edit or evaluate a choropleth map, select *Edit*, *Edit Chorpleth map*.

The interactive legend editor become visible:



Class	Symbol	[Min...	..Max>	#	# %	Average	SDCMc
1	□	7,40	10,68	2	4,08	8,55	2,64
2	░	10,68	13,96	9	18,37	12,87	3,30
3	▒	13,96	17,24	13	26,53	15,63	14,15
4	▓	17,24	20,52	18	36,73	18,63	13,00
5	█	20,52	23,80	7	14,29	22,13	8,67

Number of classes: 5

Classification method: Equal Intervals

GVF: 0,930

Figure 4.1.2
The table in the *Interactive Legend Editor* window contain for each class: symbolisation, minimum- and maximum value, number of observations (absolute and in percentage), arithmetic mean within each class and squared deviation between the observation values and the class mean (SDCMc). Large SDCMc values contribute to a poor (low) GVF value.

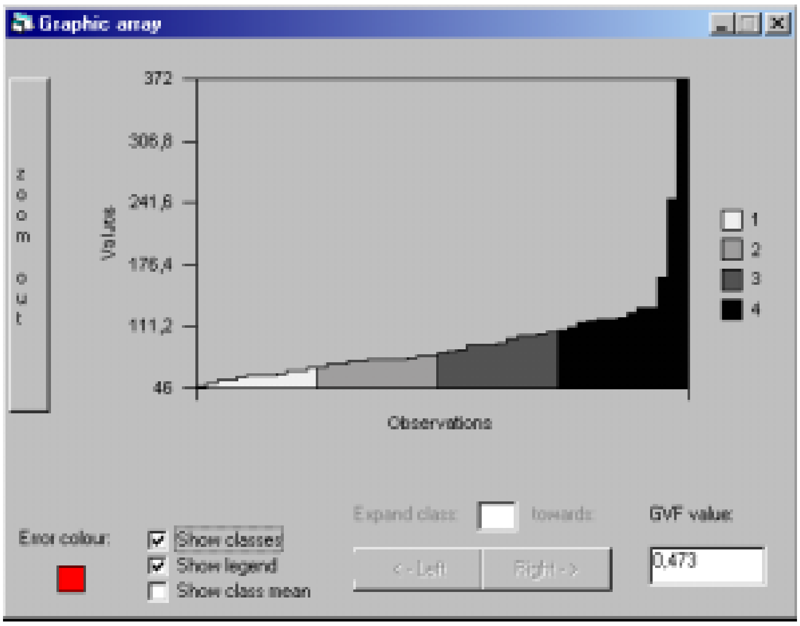
You may edit in three different ways:

- Changing the number of classes.
- Changing the method for data classification.
- Changing individual class limits by marking the '…Max>' field for the cell whose value you want to alter (the next class' minimums value will be updated accordingly).

Any types of edits made in this dialog box update the map (Figure 3.13) and the graphic array (Figure 4.2.1. ++) as these three windows are linked.

The GVF value in the lower right corner of this dialog box informs us of the "goodness" of the particular classification. In order for a classification to be considered "statistical good", the GVF value should be 0,95 or higher.

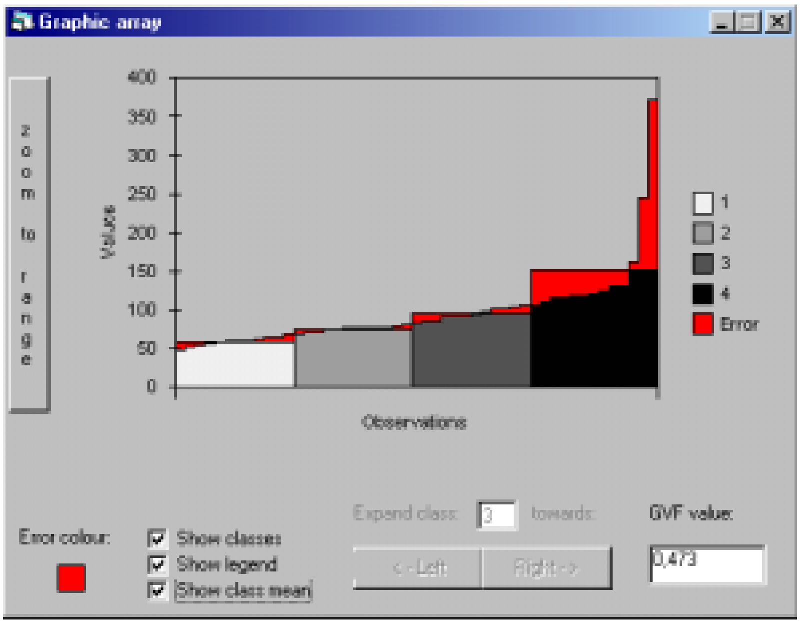
4.2. Graphic array – showing classes and class means



☑ Show classes
☑ Show legend
☐ Show class mean

GVF value: 0,473

Figure 4.2.1
The graphic array which in chapter 3 was used in order to describe the variable's distribution, might as well be used to evaluate in which way a particular classification enhance or conceal differences in the data set.
By activating *Show classes*, the observations are coloured by the actual class' symbolisation.



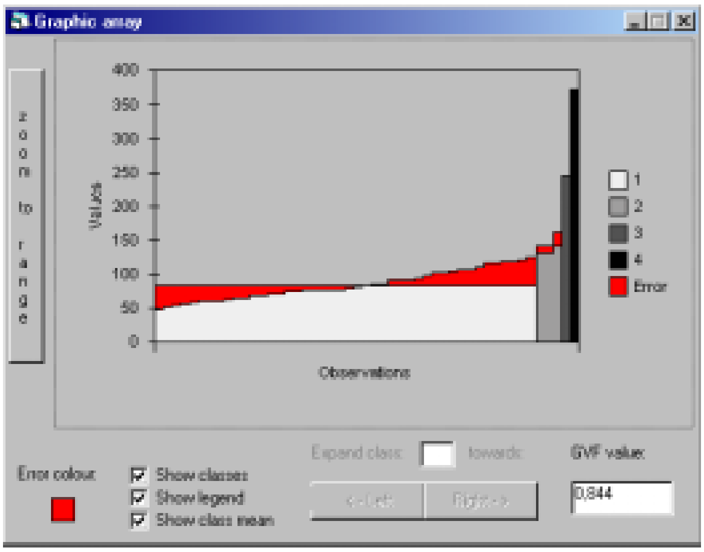
☑ Show classes
☑ Show legend
☑ Show class mean

Expand class: 3

GVF value: 0,473

Figure 4.2.2
By activating *Show class mean* the class mean is shown as a horizontal line and the deviations between observation values and the class means are coloured by red (or another "error colour" used) and are called error.

Both Figure 4.1.2 and Figure 4.2.2 show where the error is most significant for this particular variable grouped into four classes according to the quantile method. In Figure 4.1.2 it becomes evident that the fourth class has a much larger SDCMc value compared with the other classes. Corresponding, the fourth class shows larges "error area" in Figure 4.2.2 where significant variations in observation values are hidden.



☑ Show classes
☑ Show legend
☑ Show class mean

Expand class: towards:

GVF value: 0,844

Figure 4.2.3
Selected variable and number of classes is the same as in Figure 4.2.2, but the classification method is changed to *Equal Intervals*. The result is that the most significant contribution to the error is not found in the fourth class, but in the first.
As this classification results in a considerable better GVF value, it is a better classification.

4.3. Graphic array – moving observations from one class to another

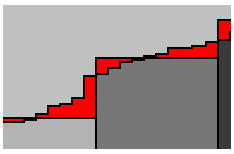
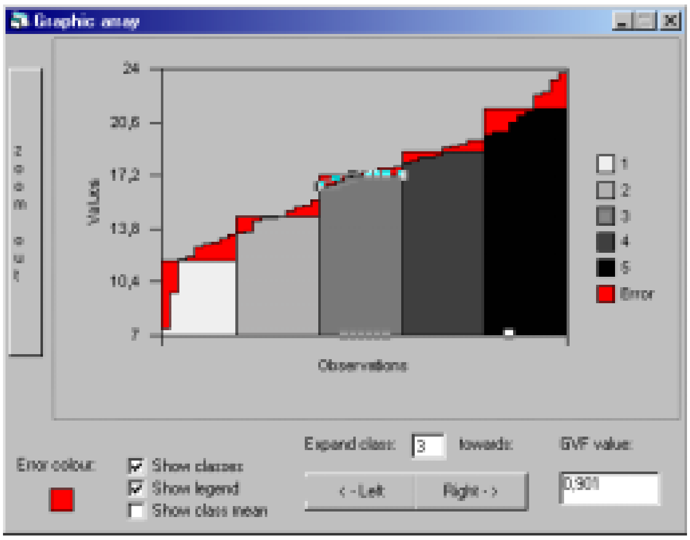


Figure 4.3.1
If a class limit is located between two observation having equal or near-equal values, the map will emphasise differences not present in the data set.



☑ Show classes
☑ Show legend
☐ Show class mean

Expand class: 3 towards:

GVF value: 0,901

Figure 4.3.2
In order to avoid that class limits are situated between equal or near-equal observation values, observations can be moved from one class to another. By selecting the class from which observations shall be moved, the command buttons *Left* and *Right* become active, which – if clicked on – move one observation to the lower or upper neighbour class respectively.

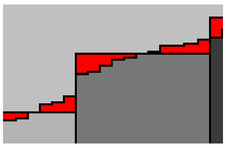
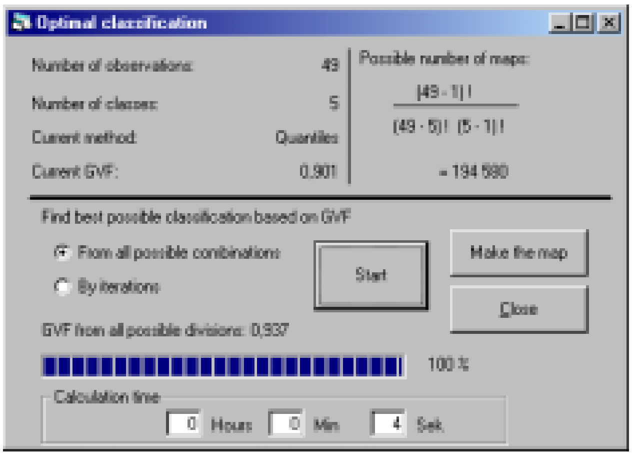


Figure 4.3.3
By locating the class limits in a manner which decreases the within class variation (homogeneity) and increases the between class variation (heterogeneity), makes the map more accurate (the GVF value increases).

4.4. Optimise data classification

When a choropleth map is generated one is able to optimise the classification (do not yield if the choropleth map is an unclassed choropleth map). Optimising data classification for the selected variable into the selected number of classes can be performed on one of two ways.

1. By evaluating all possible classification regarding their GVF values (will be very time consuming for large data set) (see Figure 4.4.1).
2. Based on the classification giving the best GVF value, observations is moved between classes in an attempt to decrease the SDCMc values and thus to increase the GVF value (see Figure 4.4.2).



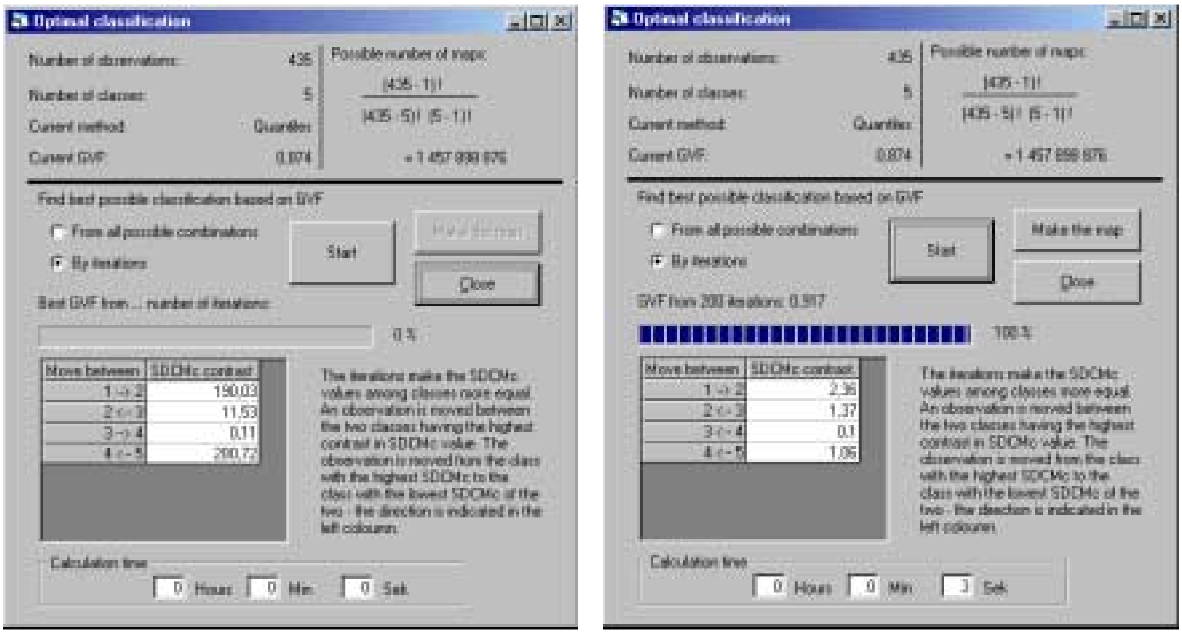
Number of observations: 49
Number of classes: 5
Current method: Quantiles
Current GVF: 0,901

Possible number of maps:

$$\frac{(49-1)!}{(49-5)!\ (5-1)!} = 194\ 580$$

Find best possible classification based on GVF
● From all possible combinations
○ By iterations
GVF from all possible divisions: 0,937
100 %
Calculation time: 0 Hours 0 Min 4 Sek

Figure 4.4.1
Finding the best possible classify-cation, based on the GVF index, from all possible combination. The variable used here is *Percentage of population above 67 years in the counties South and North Trøndelag in 1999* (see Rød, 2000), which can be classified in 194 580 different ways into five classes. The highest GVF value obtainable for this variable grouped into five classes is 0,937.



Number of observations: 435
Number of classes: 5
Current method: Quantiles
Current GVF: 0,874

Possible number of maps:

$$\frac{(435-1)!}{(435-5)!\ (5-1)!} = 1\ 457\ 898\ 876$$

Find best possible classification based on GVF
○ From all possible combinations
● By iterations
Best GVF from ... number of iterations:
0 %

Move between	SDCMc contrast
1 -> 2	190,03
2 <- 3	11,53
3 -> 4	0,11
4 <- 5	200,72

The iterations make the SDCMc values among classes more equal. An observation is moved between the two classes having the highest contrast in SDCMc value. The observation is moved from the class with the highest SDCMc to the class with the lowest SDCMc of the two - the direction is indicated in the left colourn.

Calculation time: 0 Hours 0 Min 0 Sek

Number of observations: 435
Number of classes: 5
Current method: Quantiles
Current GVF: 0,874

Possible number of maps:

$$\frac{(435-1)!}{(435-5)!\ (5-1)!} = 1\ 457\ 898\ 876$$

Find best possible classification based on GVF
○ From all possible combinations
● By iterations
GVF from 200 iterations: 0,917
100 %

Move between	SDCMc contrast
1 -> 2	2,36
2 <- 3	1,37
3 <- 4	0,1
4 <- 5	1,06

The iterations make the SDCMc values among classes more equal. An observation is moved between the two classes having the highest contrast in SDCMc value. The observation is moved from the class with the highest SDCMc to the class with the lowest SDCMc of the two - the direction is indicated in the left colourn.

Calculation time: 0 Hours 0 Min 3 Sek

Figure 4.4.2
Finding the best possible classification, based on the GVF index, by iterations. The iterations try to reduce the SDCMc values (see *Interactive Legend Editor*) by moving observation from a class having high SDCMc value to a neighbouring class having low SDCMc value. The variable used here is *Percentage of population between 0 and 20 years in Norway in 1997* (see Rød, 1999). The procedure stops (here after 200 iterations) when new iterations do not improve the GVF value more than 0,001. The iteration technique must be regarded as unaccomplished as it tends to be locked in a loop by moving the same observation between two classes.

5. Dot density map

5.1. Preparations

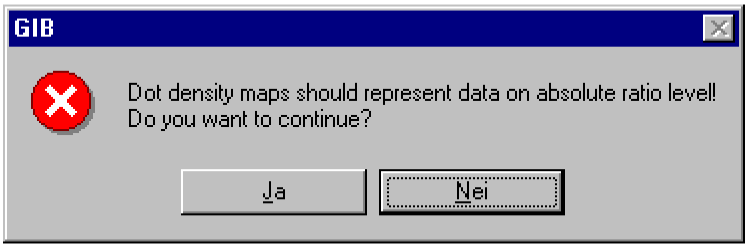
In order to produce a dot density map in GIB:

Figure 5.1.1
Select Dot density map from Map Gallery.

or

Figure 5.1.2
Click on the dot density icon.

When producing a dot density map, the variable should be on absolute ratio level. If you have selected a variable which is not on this level, you will be prompted the following message:

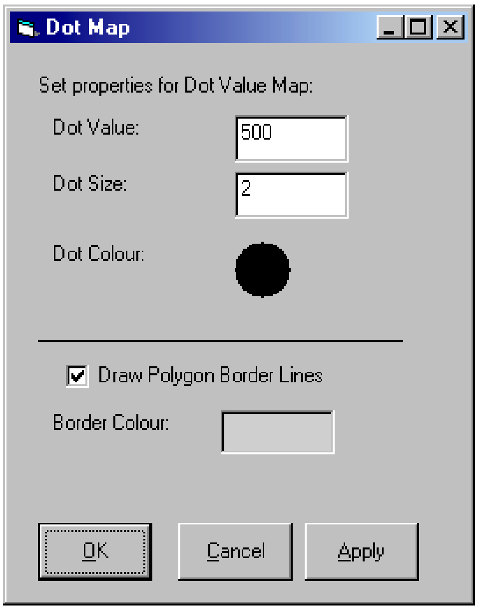

GIB

Dot density maps should represent data on absolute ratio level!
Do you want to continue?

Figure 5.1.3
Warning popping up when measurement level and representation form do not correspond.

If the measurement level is OK, the dialog box for designing dot density maps will appear (Figure 5.2).

5.2. Required options by dot density mapping


Dot Map

Set properties for Dot Value Map:

Dot Value: 500

Dot Size: 2

Dot Colour: ●

☑ Draw Polygon Border Lines

Border Colour:

Figure 5.2
Dot value is the value one dot represents, e.g. 25 men in the age 20 – 29 year.
Dot size is the physical dot size.
Dot Colour is the dot's fill colour. The dot's outline colour cannot be altered – it is always black.
Draw Polygon Border Lines
If activated polygon–borders (e.g. municipality borders) will be shown, otherwise if not.
Border Colour controls the colours of the borderlines. Click on the square to edit the colour (see Figure 3.10.2).

Click *Apply* to update the map without closing the dialog box.
Click *OK* to update the map and closing the dialog box.

5.3. The result: a dot density map

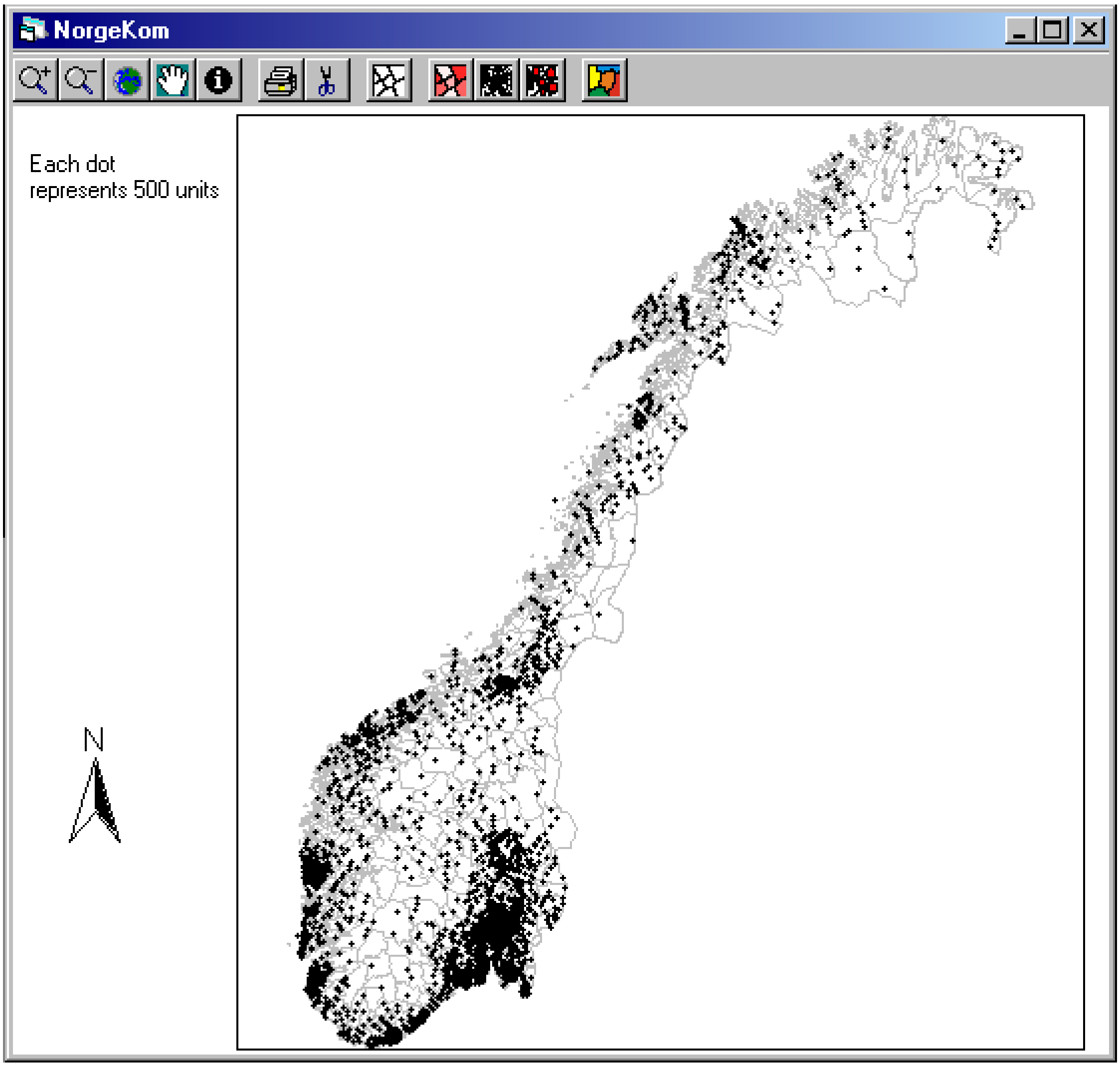


Figure 5.3
Dot density map for the variable: WM0_20_97r (number of population between 0 and 20 years in Norway in 1997).

6. Proportional point symbol map

6.1. Preparations

In order to produce a proportional symbol map in GIB:

Figure 6.1.1
Select Univariate Symbol Map from Map Gallery.

or

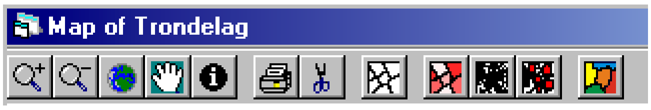
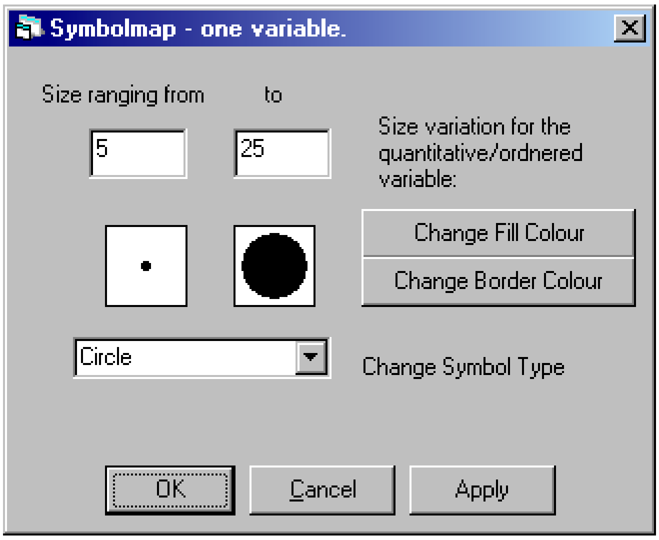


Figure 6.1.2
Click on the proportional symbol map icon

Proportional symbol maps can be used for variables both on relative ratio and absolute ratio level.

6.2. Required options by proportional point symbol mapping



Size ranging from 5 to 25

Size variation for the quantitative/ordnered variable:

Circle

Figure 6.2
The size of the proportional point symbol might be altered by entering new values for these (i.e. the fields where the values ‘5’ and ‘25’ are present). The symbol’s fill colour can be changed by clicking on *Change Fill Colour* (see Figure 3.10.2 for further details). The symbol’s outline colour can be changed by clicking on *Change Border Colour* (see Figure 3.10.2 for further details). Default outline colour is white to make it easier to distinguish overlapping symbols. Four types of symbol types are available (circles, squares, triangles and crosses). Click on *Apply* to update the map without closing the dialog box. Click on *OK* to update the map and close the dialog box.

6.3. The result: a proportional point symbol map

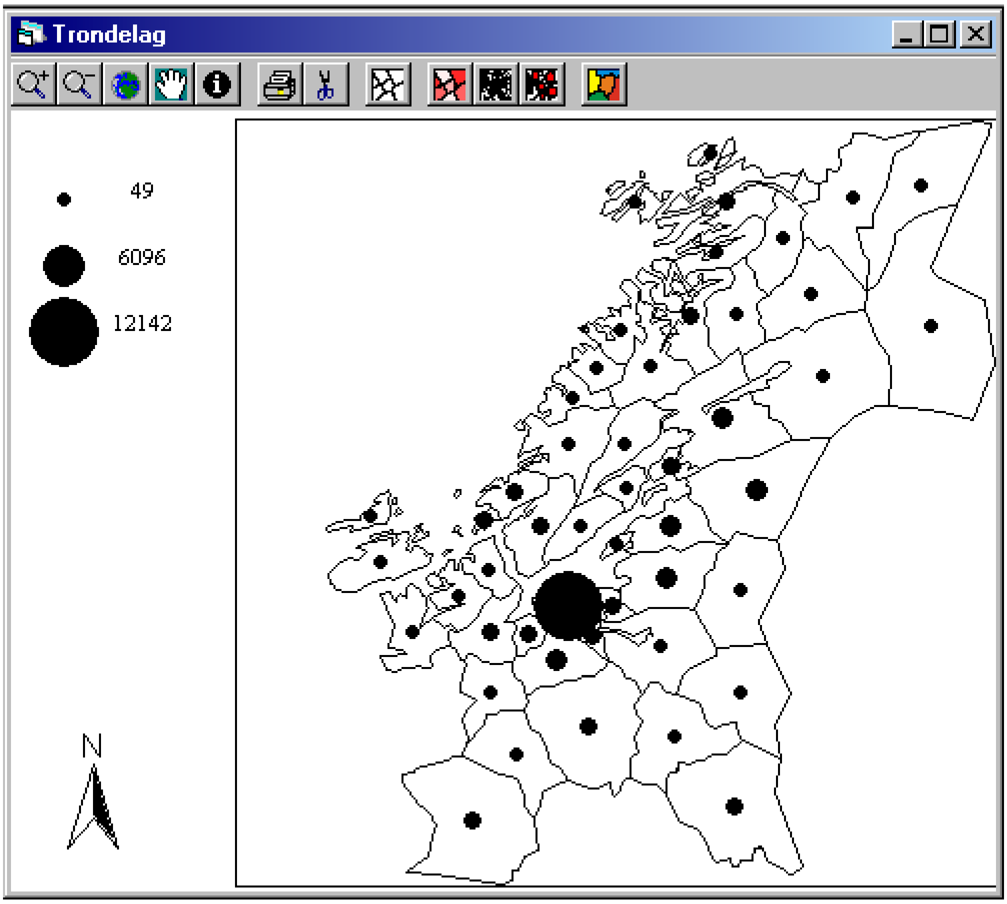
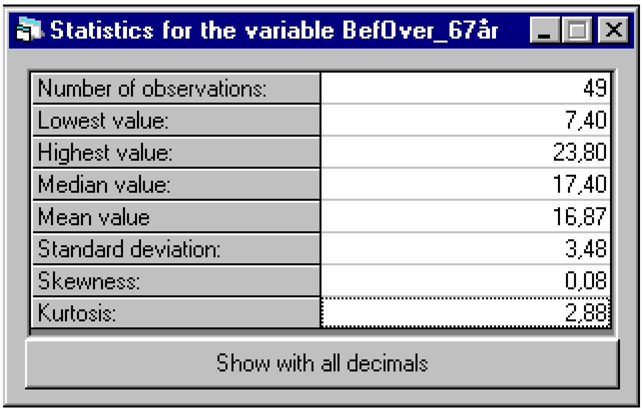


Figure 6.3
A symbol proportional maps of the variable: M20-29_93 (number of men between 20 and 29 years old in South and North Trøndelag in 1993).

7. *Additional functions*

7.1. Descriptive statistics



Statistics for the variable BefOver_67år

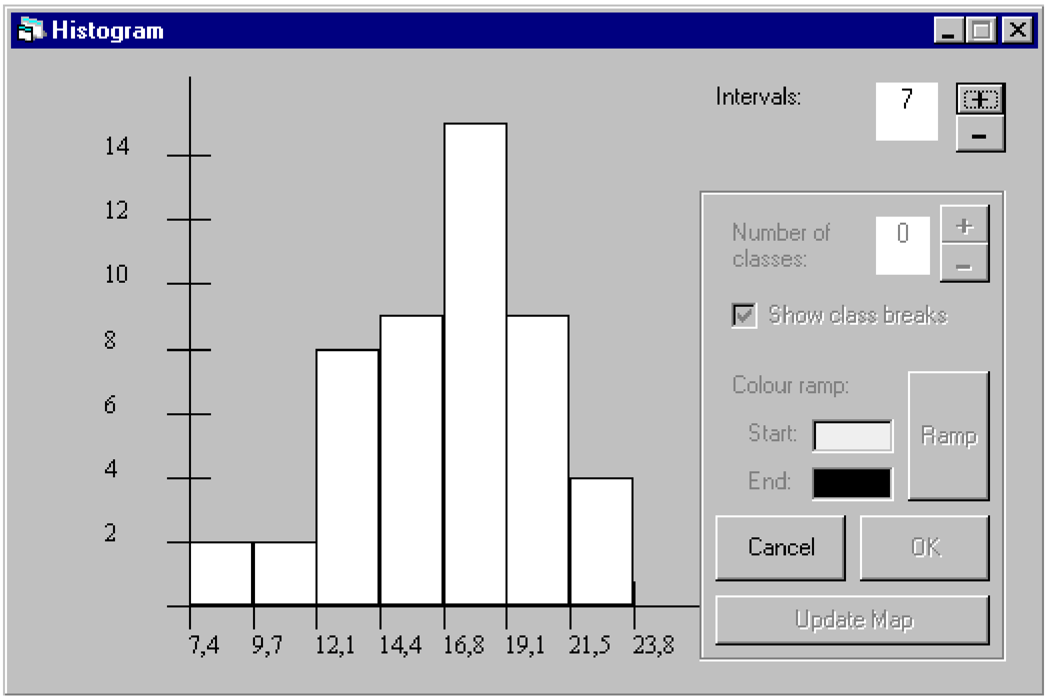
Number of observations:	49
Lowest value:	7,40
Highest value:	23,80
Median value:	17,40
Mean value	16,87
Standard deviation:	3,48
Skewness:	0,08
Kurtosis:	2,88

Figure 7.1
Descriptive statistics for the selected variable (here: PopAbove_67 , Inhabitants above 67 years in 1997 in South and North Trøndelag) is loaded by selecting *View*, *Descriptive Statistics*.

For the last two parameters:
A Gaussian distribution will have skewness equal 0 and kurtosis approximating 3. The current variable thus, approximates a Gaussian distribution.

7.2. Histogram

Histogram of the selected variable is drawn by selecting *View*, *Histogram*. The histogram is shown in one of two modus: ordinary modus or classification modus.

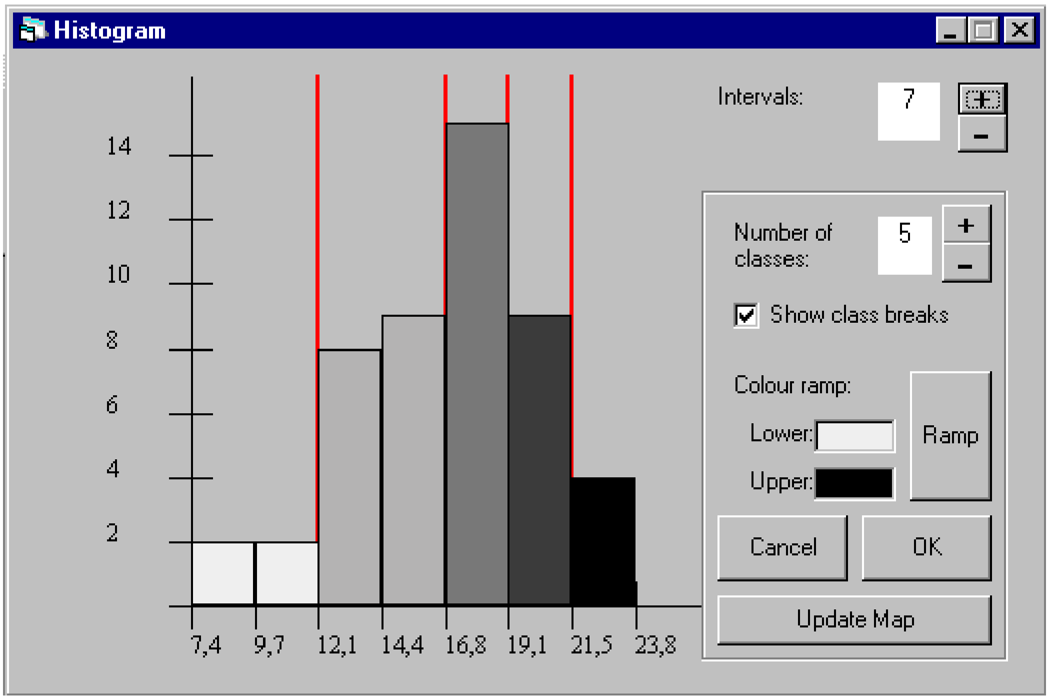


Intervals: 7

Number of classes: 0

Figure 7.2.1
Histogram in ordinary modus.

The number of intervals is by default the square root of the number of observations (here 7), but can be increased or decreased by clicking on the “+” or “-” buttons respectively. The histogram will then be updated with the given number of intervals. The intervals minimum and maximum values are indicated at the first axis. The variables minimum and maximum values are used as a point of departure for the intervals. The number of observation within each interval (with the interval range as indicated) is represented by the height of the bar and can be read from the second axis. The number of observations falling each interval is represented by the height of the bar and might be read from the second axis.



Intervals: 7

Number of classes: 5

☑ Show class breaks

Colour ramp:

Lower:

Upper:

Figure 7.2.2
Histogram in classification modus.

By selecting the classification method *natural breaks* during the choropleth mapping process (see Figure 3.4), the histograms loads in classification modus. In classification modus, the bars are coloured according to its class membership and class limits are shown as red lines. The class limits are automatically set where the difference between adjacent bars is highest. If the number of intervals or classes is changed, a new number of class breaks is determined. You may change symbolisation by clicking on the squares for lower and upper colour ramp respectively. Click on *ramp* to update the colouring of the bars.

If there are vacant positions, you may alter class limits by moving the red lines. For the red line to the left in Figure 7.2.2 there are vacant positions on both sides – hence, you may move this class border towards both left and right. For the red line to the right in Figure 7.2.2, however, there are no vacant position neither to the left nor to the right – hence, the class limit cannot be moved. By placing the cursor above the red lines, an icon is shown telling whether the limit can be change and in which direction.

⟷	Class limit can be moved in both directions.
←	Class limit can be moved to left.
→	Class limit can be moved to right.
⊘	Class limit cannot be moved.

7.3. Functionality related to the map window

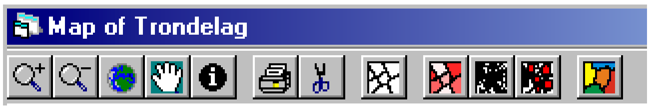


Figure 7.3
Tool menu in map window.

The tool menu consist of the following functions:

Zoom in. Click on this icon and the cursor is changed to a magnifying glass. Click on the map and it will be zoomed in.

Zoom out. Click on this icon and the cursor is changed to a magnifying glass. Click on the map and it will be zoomed out.

Zoom to full extension.

Pan. Click on this icon and the cursor is changed to the palm of a hand. By moving the cursor, you move the map.

Identify associated with a particular unit (e.g. a municipality) (see section 7.4).

Printing the map (not a layout function).

Copy the map to clipboard.

Turn of all symbolisation and regenerate the base map.

Generate a choropleth map.

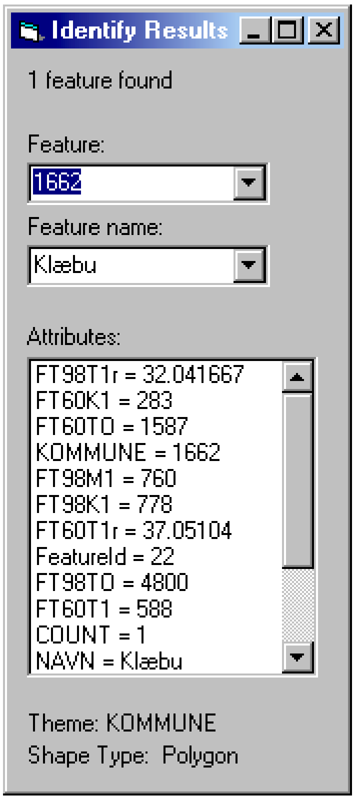
Generate a dot density map.

Generate a point proportional symbol map.

Generate a nominal area map (chorochromatic map).

7.4. Identify

ℹ



1 feature found

Feature: 1662

Feature name: Klæbu

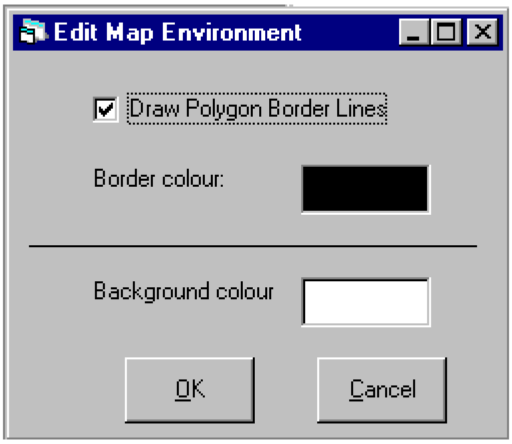
Attributes:

FT98T1r = 32.041667
FT60K1 = 283
FT60TO = 1587
KOMMUNE = 1662
FT98M1 = 760
FT98K1 = 778
FT60T1r = 37.05104
FeatureId = 22
FT98TO = 4800
FT60T1 = 588
COUNT = 1
NAVN = Klæbu

Theme: KOMMUNE
Shape Type: Polygon

Figure 7.4
By clicking on the identify icon, the cursor changes into an 'i'. By situating the cursor above one of the units in the map (e.g. the municipality "Klæbu") the information associated with this unit is shown in this window.

7.5. The map environment



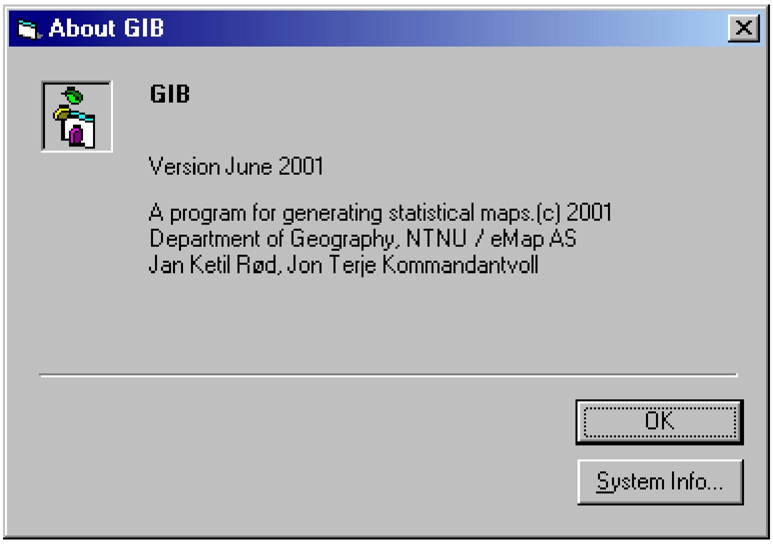
☑ Draw Polygon Border Lines

Border colour: ■

Background colour ☐

Figure 7.5
By choosing *Edit*, *Edit Map Environment* this dialog window is loaded. The commands above the line correspond with the commands under the line in Figure 5.2.
Background colour is the colour, which surrounds (behind) the map. Click on the square to change the colour (see Figure 3.10.2 for further instructions).

7.6. About GIB



GIB

Version June 2001

A program for generating statistical maps.(c) 2001
Department of Geography, NTNU / eMap AS
Jan Ketil Rød, Jon Terje Kommandantvoll

Figure 7.6
Information about GIB contain name of the originator (Jan Ketil Rød) and Jon Terje Kommandantvoll who has done some coding. Also shown is the name of the former owner (Department of Geography) and the current owner (eMap AS) of GIB.

8. How projects are saved in GIB

By selecting *File*, *Save* the project is saved with the file name given when started (see Figure 2.2) as an MS Access database with the extension *.gib under the directory <path where GIB is installed>\GIB\Data\. If you need later to load a saved project, you can do so by selecting *Open existing project* and thereafter indicate the name of the project – the project is then opened according to the saved information.

The project consists of several tables in an MS Access database. Some more background information is given below.

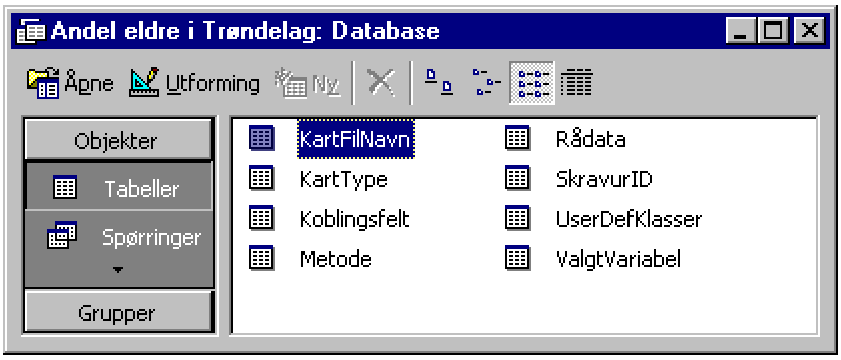


Figure 8.1
Some saved tables for the project: “Andel eldre i Trøndelag.gib” (Percentage of elderly in Trøndelag).

What is saved is (text in brackets refer to the table name – see Figure 8.1):

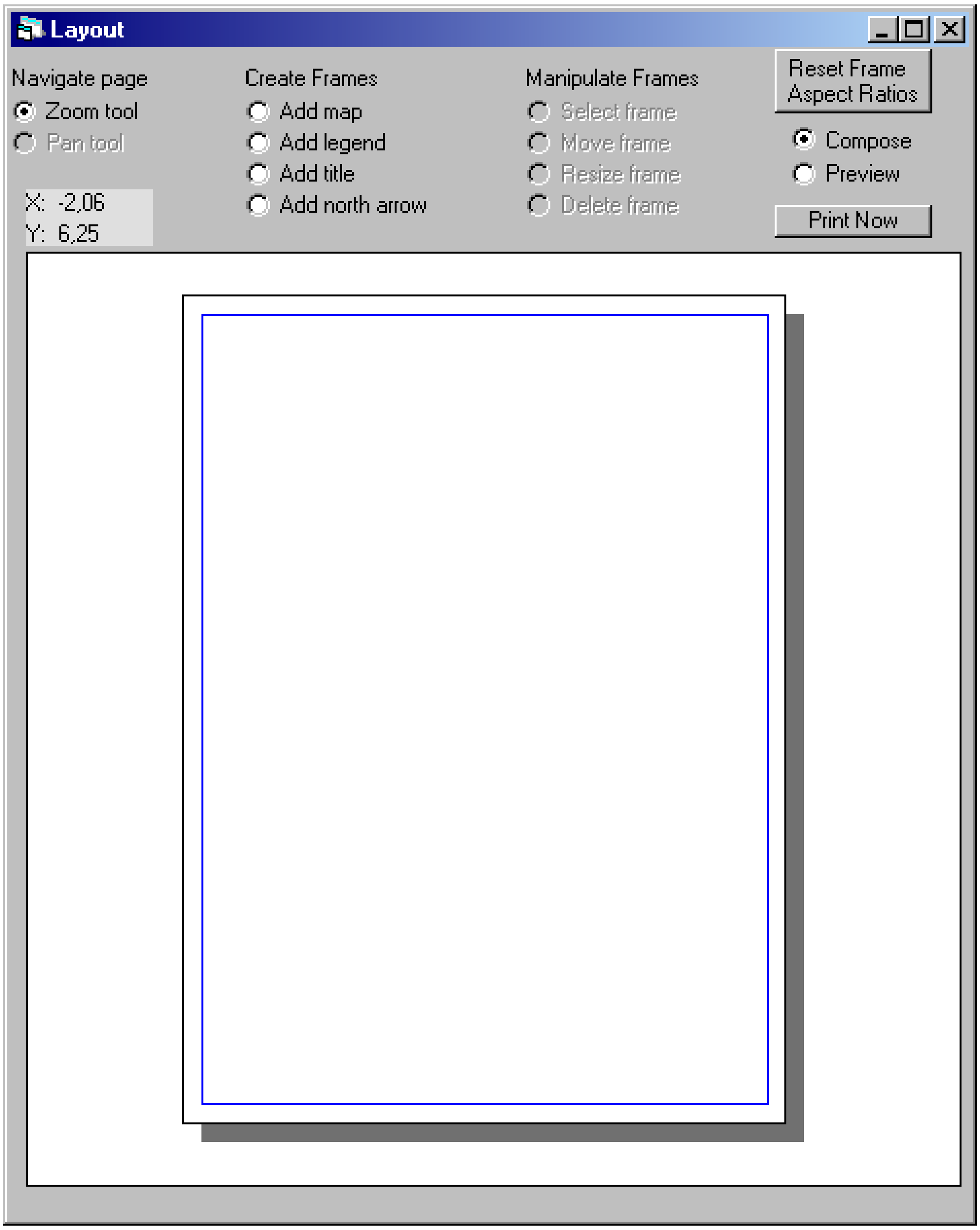
- name and location of polygon shape file and, if it exist, a point shape file (KartFilNavn)
- converted Excel sheet (Rådata)
- field used to join the map with the Excel spreadsheet (Koblingsfelt)
- selected variable to be represented in a map (ValgtVariabel)
- type of map: choropleth map, dot density map, point proportional symbol map, nominal map or simply a base map (Karttype)
- if the type of map is choropleth, the method applied for classification is saved here (Metode)
- if the type of map is choropleth, class borders and other information related to the classification corresponding with what is present in the interactive legend – see Figure 4.1 – is saved here (SkravurID)
- if class limits are user defined – see Figure 3.7 – and the user saved these, they are found here (UserDefKlasser)

Users do not need to know how this information is stored to be able to operate GIB with existing project.

9. Layout

To compose a map for printing select: *View, Layout*. The Layout window will then be visible located just right of the map window. In order to use Layout the screen resolution on your computer must allow the map window and the layout window to be located next to each other.

9.1. Layout window when loaded



Navigate page
● Zoom tool
○ Pan tool
X: -2,06
Y: 6,25

Create Frames
○ Add map
○ Add legend
○ Add title
○ Add north arrow

Manipulate Frames
○ Select frame
○ Move frame
○ Resize frame
○ Delete frame

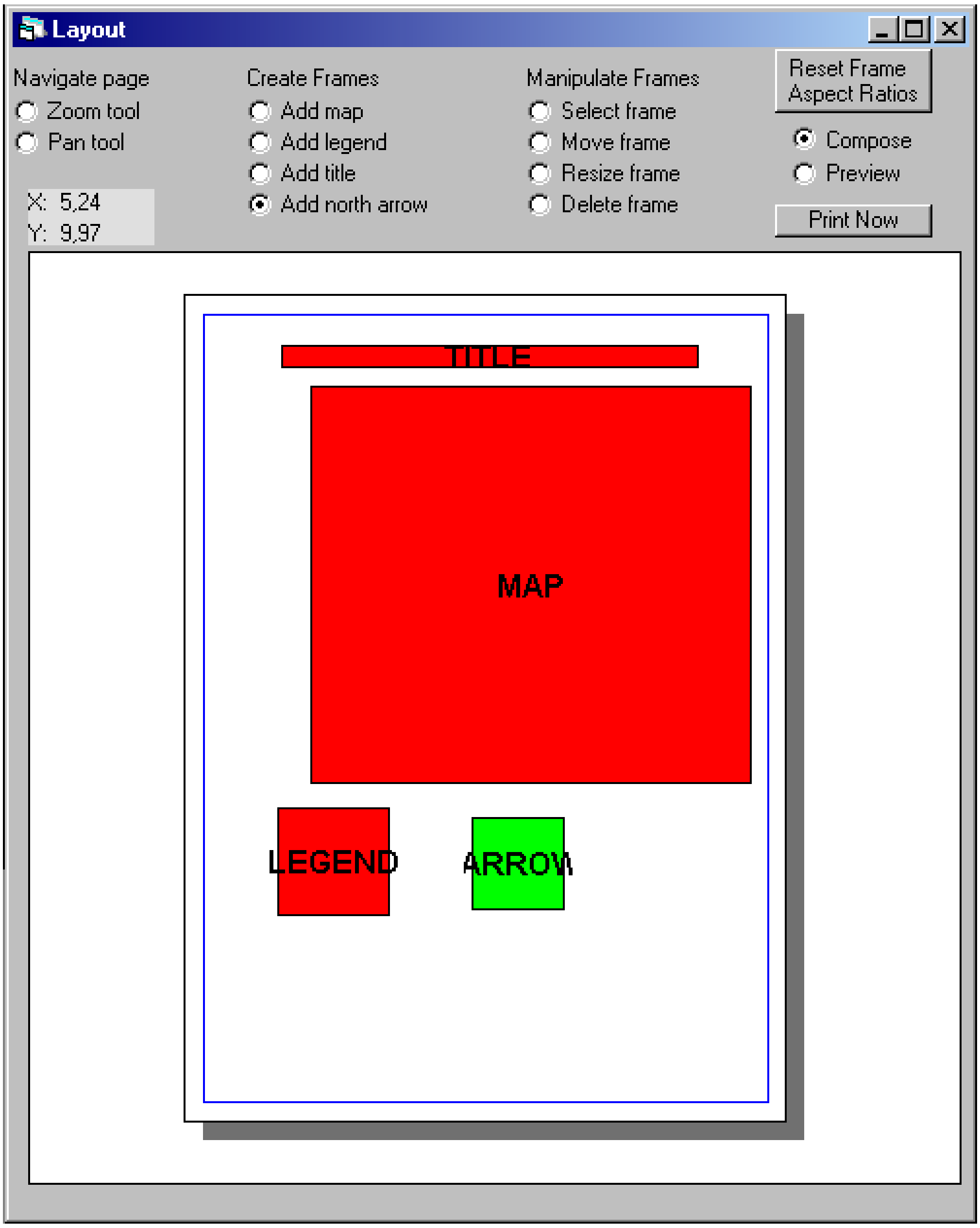
● Compose
○ Preview

Figure 9.1
Layout window in compose mode without any "objects".

The layout window is shown in one of two modes: *compose mode* and *preview mode*. When started, layout is in compose mode. Compose mode allow for locating objects "on the paper". Three types of objects can be placed: map *(add map)*, legend *(add legend)* and title *(add title)*.

Click on the option listed under *Create Frames* for the object you need to locate and “drag” a rectangle on “the paper” in order to indicate its location. Repeat for the number of objects you need to locate. The result after having located three objects on “the paper” might look like Figure 9.2.

9.2. Layout window in *compose* mode



Navigate page
○ Zoom tool
○ Pan tool

X: 5,24
Y: 9,97

Create Frames
○ Add map
○ Add legend
○ Add title
● Add north arrow

Manipulate Frames
○ Select frame
○ Move frame
○ Resize frame
○ Delete frame

● Compose
○ Preview

Figure 9.2
Layout window in compose mode with four “objects” added.

The object last added is shown with a green colour, the other in red. The object with a green colour is the selected object which means it can be moved (*move frame*), resized (*resize frame*) or deleted (*delete frame*) depending on activated option. To select another object, click first on the option *select frame* and then on the object, which then will be coloured with green.

Click on *Reset Frame Aspect Ratios* to assign the objects same relations between height and width as in the map window. The *preview* option allow you to preview the result.

9.3. Layout window in *preview* mode

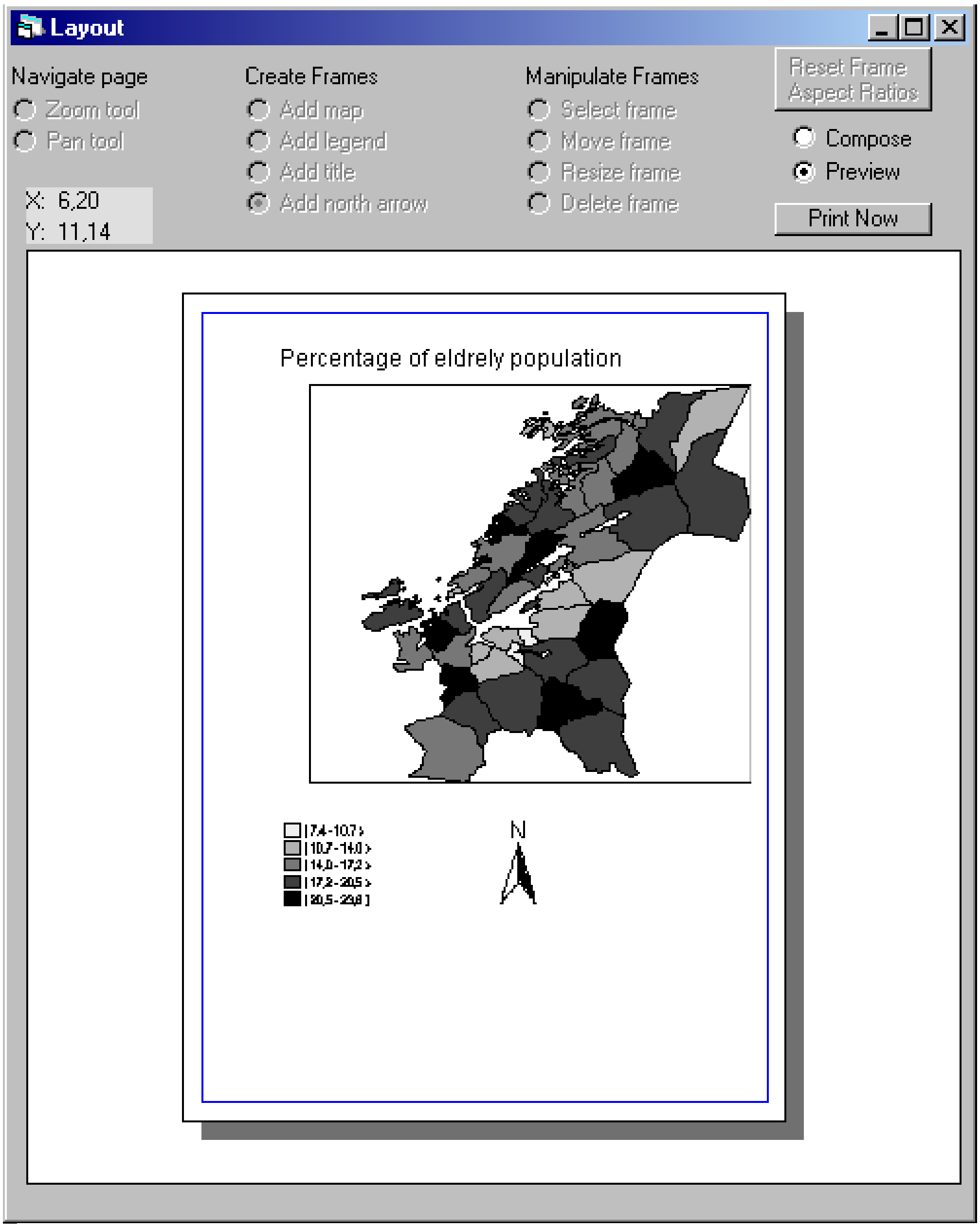


Figure 9.3
Layout window in preview mode. In preview mode none of the other options or commands are available except the print command button.

References

Cauvin, C., Reymond, H. and Serradj, A. 1987. *Discrétisation et representation cartographique*. Montpellier: GIP Reclus.

Dent, B.D. 1999. *Cartography : thematic map design*. Boston: W.C. Brown Publishers.

Robinson, A.H., Sale, R.D., Morrison, J.L., Muehrcke, P.C. 1984. *Elements of Cartography*. New York: Wiley.

Rød, J.K. 1999. *Vurdering av funksjonalitet ved fremstilling av koropletkart I Excel*. Kart og Plan. 59 (1): 36-42.

Rød, J.K. VI. *The selection of class intervals revisited*. Submitted to Transactions of the Institute of British Geographers.

Slocum, T.A. 1999. *Thematic cartography and visualization*. Upper Saddle River, N.J. : Prentice Hall.

Appendix 1: Preparing the Excel file

If GIB shall be able to read Excel files and connect it with a base map, the following conditions must be fulfilled:

1. The statistical data must be stored in the first Excel sheet.
2. The first column in this sheet must contain connection field. If you are using base map on municipality level, the municipality numbers will likely be used as connection field (i.e. 1601 for the Trondheim municipality in Norway).
3. Column heading must exist for all columns (i.e. “Kommune”, “Navn”, “BefOver_67år”, …).
4. Column headings cannot exceed the upper row. Second row on must be filled with data.
5. Diagram or other objects must not be in this cheet.
6. No open rows.

KOMMUNE	Navn	BefOver_67år	…
1601	Trondheim	12,6	…
1612	Hemne	14,6	…
1613	Snillfjord	21,3	…
1617	Hitra	18,0	…
1620	Frøya	19,4	…
1621	Ørland	13,2	…
1622	Agdenes	18,4	…
1624	Rissa	17,5	…
…	…	…	…

Nice!

~~Befolkningsdata 1997~~

~~Navn~~		~~BefOver_67år~~	~~…~~
~~Trondheim~~	~~1601~~	~~12,6~~	~~…~~
~~Hemne~~	~~1612~~	~~14,6~~	~~…~~
~~…~~	~~…~~	~~…~~	~~…~~

Not nice!

Appendix 2: The shape map file format

A shape file is an ESRI (Environment Systems Research Institute) data file format for storing geographical data in vector format (i.e. point, line and polygon representation).

A shape file is basically a collection of (at least) three files. All files have same name but have unique extensions. The three basis files are:

main file:	*.shp	holds the shape's geometry
index file:	*.shx	holds indexes for the phenomena's geometry
dBase table	*.dbf	holds a table with attribute information for each shape

In GIB, two shape types are used: polygons and points. The polygons are likely to be borders for administrative units, for instance country, county, municipality and the like.