ELECTRONIC STATISTICAL ATLAS PRODUCTION IN TURKEY: POTENTIAL AND IMPEDIMENTS

N. ULUGTEKIN[1], I.O. BILDIRICI[2], O.F. GOKCEN[1]

[1]ITU Faculty of Civil Engineering, Dept. of Geodesy & Photogrammetry, Div. of Cartography, 80626 Maslak ISTANBUL TURKEY
[2]Selcuk University, Faculty of Engineering, Dept. of Geodesy & Photogrammetry, Div. of Cartography, 42079 KONYA TURKEY

In this paper, an electronic statistical atlas as a case study is described. The authors have completed development of a prototype of the atlas, which present the spatial information in the statistical yearbook in an atlas structure that contains thematic and animated maps. The paper also reports how statistic geometric and non-geometric data are captured and combined. Aim of this research is to show the possibilities of producing such an electronic atlas and bring suggestions for data modeling and design problems on a prototype, which is produced and will be produced by authors. The potential of statistical data of spatial character in Turkey is evaluated. The problems related to spatial and non-spatial data capture; update problems are also in detail discussed.

1. Introduction

Development of Electronic Atlases is a research field in cartography. Some developed countries have created their own electronic atlases or national atlas information systems however [Ormeling 1995]. In Turkey, there are several atlases of different design and quality. Unfortunately there is no high quality atlas produced by an experienced team for certain group of users. Especially school atlases are full of basic mistakes that lead misunderstandings [Uluğtekin 1989]. In Turkey there is also no also National Statistical Atlas. Some of the atlases contain maps based on statistical data. It is possible to find electronic media examples with various aims on market but they are generally commercial products, which are not produced by experts.

Atlases are systematic and coherent collections of geographical data in analog or digital form, presenting a specific area and/or one or more geographical themes, based on a specific objective or narrative, together with tools for navigation, information retrieval, analysis and presentation. *Atlas structure* refers to the basic framework according to which the maps have been incorporated in the atlas. It is based on the emphasis on specific areas or map themes and on the sequence in which these maps follow each other. The *narrative* is the term for the objective behind the atlas. Map sequence and map emphasis contributes to the atlas narrative. The maps on an atlas spread together tell a story. Basically, atlas structure is a matter of emphasis and sequence. Emphasis is the number of maps or the number of map pages dedicated to a particular area or subject. Sequence refers to the order in which the maps are presented [Ormeling 1997].

In a digital atlas data that answers specific requirements, as it should emphasize specific trends. Animation is other super component of electronic atlases [Peterson 1995, Peterson 2000]. Development of electronic atlases is so important because of their superior storage and analytical capacities. They have the ability to provide more, different views of the same data, by allowing for changes in definition, in classification or categorisation, in symbolisation or by providing different color schemes [Ormeling 1997]. "*The creative behind the majority of such atlases belongs to scientific societies and organisations which felt that the preparation of atlases was a task of great scientific value, prestige and national importance*" [Siekierska 1993]. In order to believe the same task that pilot project is made by small an academic group and budget for researching, developing and learning.

One of the main information sources of an atlas is statistical data. Census data is available mapping for fast, accurate, understandable maps and also dissemination of spatial statistics data. Themes represented in socioeconomic maps are often derived from statistics related to topics such as census data, infrastructure, housing, employment, trade and agriculture. Visualization of statistical data is various types of graphs. It is possible to represent numerical statistical data with pie graphs, bar graphs or graduated symbols. It is also possible to convert the data into choropleth map or dot maps, through which recognition of the spatial trends would be obtained.

2. Aim of the Study

The aim of the study was to present the statistical data based on the provinces of our country in an electronic atlas medium. This brought up the need to obtain geometric and non-geometric data in provincial and later on in district level. In here the geometric data that are being used are in point (for city centres) and area characters. The non-geometric data such as population growth and gross national income per person are in statistical character. The main goals of this project is to offer end users the statistical data that reflects the reality of the country within the rules of cartographic communication, to obtain the basis of an electronic atlas and to create a prototype product. Such an atlas can be used in education as well as significantly assisting a range of user groups to recognise the realities of our country better [Ulugtekin *et al*. 2000].

In this study, a map that covers the whole country in the same scale is used as a geographical base in order to be used in spatial comparisons. Spatial statistical data along with concerning subjects and census data of different dates is used for comparison. The visualised data within the electronic atlas enables a general view of Turkey, with the help of thematic maps drawn to a scale of 1:8 000 000. Thus the map data, enriched with visual and animated supports, is presented to the user in the most suitable, quickest and cheapest manner. The users will only be capable of viewing thematic maps that are based on certain questions with previously prepared scenarios [Ulugtekin & Bildirici 1999, Ulugtekin *et al*. 2000].

In the first versions of the study, the need to use PowerPoint VIEWER, freeware Microsoft software was necessary for the users to reach the atlas data. The main functions are presentation of the map and statistical data according to the subject. This version is included some animations. In the latest version of the study, GLPRO software has been used to remove the need of supplementary software, thus enabling the product to be directly used in a CD-ROM format.

3. Data Capture

Statistical maps are used to establish the location of many types of boundaries, to prevent omission or duplication of information, to aid enumerator canvassing, to facilitate coordination between various offices and between office and field staff, to settle administrative jurisdictional problems, to identify segments used in area sampling, to help determine the number and distribution of enumerators and supervisors to be employed and to foster comparability of data from census to census [Gorden 1987]. Statistical surveyors collect socioeconomic data, either continuously or in censuses. Raw data sets need to be referenced. Only after the data sets are properly referenced can interpretation and representation occur [Frappier & Williams 1999]. The spatial databases contain geometrical and non-geometrical data. The geometrical database provides base elements and references for statistical data. These can be locations linked to administrative units or the administrative units themselves. The non-geometric database contains all statistical data as well as meta data about each set of data. This meta information defines the spatial level referred to, the time dimension, the source, as well as the way of operation, and the type of data which determines the type of representation for the map [Lambrecht & Tzschaschel 1999].

As mentioned before, the creation of homogeneous statistical units is very important during the preparation of these types of maps. The change in fields while collecting data from census to census creates big problems in mapping out the comparative elements and preparation of animated maps based on time. We can even say that it makes it impossible to prepare such maps. In Turkey the administrative boundaries change frequently in province basis and the census is done taking into account these administrative boundaries. A newly formed province is created from districts separated from another province or other provinces. The administrative boundaries are not definite and permanent by natural objects. For this reason the State Institute of Statistics (SIS) yearbook lose their validity in a short period of time and as the data within the yearbook cannot be manipulated easily to cover for the new administrative boundaries, the data are printed according to different administrative units separations within the same yearbook. As different administrative units' separations form the basis for publications of yearbook from year to year, it becomes impossible to do time comparisons for most topics elements. Existing statistical data is sufficient for years 1980, 1985, 1990, 1995, 1997 statistical area boundaries (administrative boundaries) are quite different. As mentioned before this is an important handicap to interpret the data. Therefore all geometric and non-geometric data should be updated every year, in some cases every three months. In Turkey, this update process has been hardly realised because of the lack of communication between governmental organisations [Ulugtekin & Bildirici 1999].

3.1 Geometrical Data Capture

It was found to be suitable to collect the spatial data that will form the geometric base of the electronic atlas firstly in provincial level (then in district level) and thus the scale of 1:1 000 000 was selected. To assist the users in recognising the geographical locations better, lakes, dams and rivers that are meaningful to show in 1:1 000 000 scale are digitised, along with the administrative boundaries, to the maps that are going to be used in the atlas. Apart from the administrative boundaries, all city and town centres of the administrative regions are digitised as point objects. Upon this digitalisation, cartographic objects in 25 different layers are obtained [Ulugtekin & Bildirici 1999, Ulugtekin *et al*. 2000].

The Political Map of Turkey that is used as a base, is created by General Command of Mapping using two standard parallels in Lambert conform conic projection. The data is converted to UTM projection, as it would be easier to add data from other sources if needed. Another reason for this conversion is that UTM is an international standard and is easily defined in many GIS software. But Turkey falls into 4 different zone in UTM system. In the produced atlas, as 1:8 000 000 scale will generally be used in the maps that will show the whole country, in order to avoid discontinuity all data is converted into zone 36 (λ_0=33°). The maps obtained in this manner are observed to be in acceptable deformation limits for the scale used. The scales of screen maps is affected by the monitor being used, dimensions of the monitor and even the colour choices (256 color, 16 Million color) that are being used, it should not be seen as a definite value as in analog maps. The most important aspect that determines the scale is the pixel size of the monitor being used. One pixel is 0.28 mm for standard monitors and goes down to 0.21 mm in top grade monitors. Due to this reason, the physical dimension of the monitor size varies by small amounts.

The Political Map of Turkey scaled 1:1 000 000 was digitised in AutoCAD Map environment. After obtaining the point and linear data, the administrative regions (provinces) that are to be linked with the non-geometrical data are determined by polygon topology. In order to eliminate digitalisation errors, line-cleaning actions are applied to data before the polygon topology is created. The polygon topologies for provinces and districts were built in the same environment. The boundaries of provinces were simplified in order to have a sufficient generalisation degree for the 1:8 000 000 scaled background maps for the thematic maps of whole Turkey. The line simplification tool of AutoCAD Map was used. Although the algorithm of this tool is not exactly known, sufficient results were obtained, after tests with various tolerance values. After the provinces were created as closed polylines in AutoCAD environment, they were linked to a table that contained license plate numbers and names of provinces. These plate numbers here were thought as a key field for making new links with different non-geometric data. Such a key filed was also necessary in MAPINFO environment.

The biggest problem in obtaining and determining base map data is the unbelievably fast rate of change in administrative boundaries. All the base map data that was obtained with so much difficulty and effort lost its validity with the decision of making Duzce, where is a district before, a province at the completion stage of this project. On the other hand, the non-geometric data prepared in administrative units has to be updated again according to the new divisions after every change. In this concept, the reality of the contradiction between the number of districts in the Turkish Political City Map and the non-geometric data sources can be seen as the result of this rapid change of administrative units.

3.2 Non-Geometrical Data Capture

In Turkey, the statistical information is collected and published annually in form of a yearbook by State Institute of Statistics. Recently, statistical data has been made available in digital form. This digital data can also be viewed on the Internet, but some of the data tables can not be downloaded as text, because they are published as image files on the Internet. The statistical yearbook contains large amount of tabular data that has spatial character with a few thematic maps, which are poor products in cartographic sense [Ulugtekin & Bildirici 1999, Ulugtekin *et al*. 2000].

The data in the table are listed according to the license plate number of provinces and data regarding the districts are listed within each province in a particular standard. These data are set as integer numbers and can be seen as a key field that identifies solely the particular administrative unit concerned. The tables that form the basis of relational database in EXCEL software have been transferred to MAPINFO software. As pointed out before, the rapid change in administrative divisions causes the geometric base data and also the non-geometric data to lose their validity rapidly. It is much harder to update non-geometric data compared to geometric base data. After an administrative boundary change, one must go back to the census data collected to process the data (such as population, migration, etc.) according to the new administrative units. In order to have a meaningful comparison this process must be done with at least two censuses back. No need to say that these kinds of processes are time consuming and require a lot of work.

4. Production Strategy and Software

It is not possible to use single software or a software package in electronic atlas production field. Due to this reason various software with different functions have been used within this project. The use of different software causes the data processing time to increase in atlas production, map design and also difficulty of updating maps.

Furthermore, data conversion between different software medium can also be a problem. Many experts specialised in atlas production point out this problem. A way to overcome this problem is to develop software unique to the project [Bar & Sieber 1999, Schneider 1999, Smith 1999]. But even this is not a very efficient solution.

A study on in State Institute of Statistics (SIS) yearbook has revealed that data of spatial character for all subjects are not present. Data regarding many subjects is for the whole country and not suitable for making thematic maps in provincial or district basis. Another problem is data being too old soon and not being in suitable for present day's administrative divisions. Difficulty in obtaining data valid for today is supplemented by frequent changes in the administrative divisions. After this study, taking into account the data published SIS web site main topics that are to be covered are chosen as follows: Turkey on the World, Turkey from Space, Turkey Administrative, Demography: population growth, population density etc., Vital Statistics: migration, marriages and divorces etc., Economy: expenditure on the gross domestic product etc., Transportation, Health, Education, Environment.

In order to assist the users in comparison, a base map has been created at the same scale for every subject in the same generalisation level. The maps produced here are screen maps. The maps designed for 800X600 pixel monitor dimensions are approximately at a scale of 1:8 000 000. The thematic maps produced in the MAPINFO environment were then converted into a standard picture file format that can be easily used in other software environments. As a standard picture file format Windows Meta File (WMF) was selected. The WMF format is a vector file format, from which any standard rater image format like BMP can be obtained.

Another advantage of using WMF files for the thematic maps within the atlas is that it enables aesthetic improvements and arrangements to be made using graphic design programs like COREL DRAW, ADOBE ILLUSTRATOR, etc. The maps, which are prepared in MAPINFO environment, require a supplementary arrangement, as it is very hard to combine the map and the legend together. Due to this problem maps and legends are prepared as separate WMF files and combined using COREL DRAW. Arranging the frame of the map and other similar jobs are also done in COREL DRAW software.

The animated maps used in the atlas are prepared with the 3D Studio MAX software. This software gives the users the possibility to create animations in any known file format. It has a cartesian coordinate system likes AutoCAD but do not recognise map projections. Map scale can not be defined directly. It has no direct access to any database environment. Manual editing is required if any animated map should be updated.

The base geometric data used in animated maps are imported from AutoCAD. Autodesk produces both software, because of that there was no problem in data flow. The prepared animation files are converted into a standard AVI format and transferred to multimedia production software (GLPRO). Since the AVI format is supported by

WINDOWS 95/98/NT operating systems, it can be used in every WINDOWS platform or in the Internet. The disadvantage of AVI format is that it takes up a large disk space. Due to this restriction, even though the static maps of the atlas are prepared at 800X600 pixel dimension, the animated maps have a dimension of 320X240 pixel. This restriction in image sizes means that the contents of these animated maps are also restricted compared with static maps.

It is possible to prepare animation by using PowerPoint. Animation about the change in province boundaries is prepared. This kind of animation does not have a size restriction like AVI files. Another advantage of using PowerPoint is that the prepared presentation can be transferred to the Internet medium. A plug-in must be added to the INTERNET EXPLORER web browser to do this. This kind of application can only be done with INTERNET EXPLORER and NETSCAPE cannot be used.

Electronic atlases can be distributed with CD-ROM or via Internet. These two methods have advantages and disadvantages. Using CD-ROM media, one has good archiving opportunities but updating problems. Huge amount of data can be easily distributed. The usage of data on CD-ROM can no be easily restricted. CD-ROM media can be illegal copied. In comparison with Internet, production and distribution of CD-ROM media are more expensive. In the Internet publication, the down load time of high quality pictures and animations can cause problems according to speed of connection. In order to attract the visitor download time should be minimised where possible and no additional software should be required.

GLPRO is a fast and flexible multimedia preparation programming language. GLPRO can be used in preparing and developing multimedia applications, presentations, data guides, introductory discs, high performance music CD's, interactive games and educational software, enabling them to work in fastest form with using the least amount of memory. GLPRO presents multimedia developers to work in a cross software development medium that can be used in conversions between different operating systems and different hardware. Today there is only compiler for MICROSOFT WINDOWS environment. It is said that compilers for MACINTOSH and LINUX systems are to be finished quite soon. With GLPRO software, latest technologies used in multimedia can be applied. A multimedia application created with GLPRO consists of only one file, which does not require an installation of particular software, and no DLL files other than the standard operating system DLL files are needed. This is a great advantage on behalf of the software. GLPRO software is a freeware program as long as requirement of not using it commercially is met.

5. Design Problems of Screen Maps

Static map making software also have cartographic design problems: producing maps according to minimum display configuration, choosing cartographic symbol sizes on the screen map, difficulties of synthetic or complex map making, etc. The problems also present in the multimedia product design and in the Internet applications. In order not to want to use scroll bars, the graphic map is smaller than the screen size. As a result, the maps have to be reduced in size and therefore generalised in detail, and other minimum dimensions also have to be applied [Lambrecht & Tzschaschel 1999]. There will also be the possibility of various clicking boxes or clicking links around a map, which then lead to other maps, to further information [Ulugtekin & Bildirici 1999].

Due to the physical structure of the monitors, the pixels are in the shape of squares. Due to this reason the smallest readable diagonal text differs from vertical and horizontal text which can be much smaller. When horizontal, vertical and diagonal text is written with the same text size, the sharpness of the diagonal text is much worse. The effect of pixels can be seen much easier in diagonal text. The same is true for diagonal lines.

The scale seen on the monitor will change according to the physical dimensions of the monitor. So it is meaningful to use graphic scales in screen maps. A numerical value such as 1:8 000 000 would not be useful to the user. It would only give a rough idea.

The minimum standard for colour must be 16 million (24 bit). But the physical appearance of a colour on the monitor may vary with regard to graphic card and monitor being used. The value and saturation changes of design colours in user's hardware have been seen quite frequently. So we can deduct that we have to choose the least amount of classes possible when working with choropleth maps.

Due to these restrictions coming from the monitor, the base maps of thematic maps must be more generalised than those being used for classic map studies. In this work, the names of the provinces are not written down to avoid confusion with pie and bar charts. The restrictions due to monitors also make it more difficult to present more than two subjects at the same time in thematic maps. If a lot of map elements are tried to be shown on the monitor at the same time, the users will have difficulty in understanding the information presented through the map [Ulugtekin e*t al.* 2000].

Several mapping programs are available for personal or microcomputers that create choropleth maps or point symbol maps, but the user must be aware that if raw data are entered into these programs, misleading maps may be produced. Unless the program can compute areas of enumeration units and calculate densities from these, the error just described will result. Therefore, when using such programs, the data must be entered as ratios, densities, or percentages. On statistical maps, cartographers use a variety of quantitative symbols to depict magnitude information. The visual variables, which evoke the proper image of quantitative symbols, are size, pattern texture, colour value, and colour intensity. These variables give an impression of magnitude rather than simply location [Tyner 1992, Ulugtekin & Bildirici 1999].

A major incentive for the atlas is the idea, to visualise complex and spatially differentiated information and to present it in an understandable manner, also for readers with relatively little experience in using thematic maps. For the cartographic part, special attention has to be paid to the legibility of the legends. This does not mean, that complex representation and the use of abstract symbols and indicators have to be avoided, but an effort will be made to explain such complexities in special boxes and by specially designed legends [Lambrecht & Tzschaschel 1999].

6. Conclusion and Recommendations

For the public to learn and understand geographic data about the terrestrial small and large areas, “correct” maps are of utmost importance. The data/information and cartographic quality of atlases especially designed for schools must be inspected before being used. As the controlling and auditing body of all atlas as educational material is the Ministry of Education, a control system responsible for this quality control can formed by this ministry and SIS working together to assist in supporting map use and production of “correct” maps will be very useful.

The main problem the developing countries face in trying to create census maps is the lack of foundation base maps. The statistical borders must be set rigidly in order to evade any repetitions or omissions. Many geographic problems such as migrant population, accessibility and border changes from one census to another must be overcome. As a result; for countries that has not invested in thematic map creation, it is strongly advised that they use the rich data source of a census to create their thematic maps as soon as possible. This can only be done with cooperation of experts in various disciplines to work together in creating a National Statistical Atlas.

The experience and know-how obtained from this project will assist similar studies that are going to be done in this field. We have summarised below the experience we have gained from this project and the study fields we aim to explore in the future. The aim of this Istanbul Technical University Research Fund supported project, which had a very limited time, and budget was to show that these kinds of studies can be realised and to share the data and findings of this project with other users through the Internet and CD-ROMs. The researchers of this project are trying to further develop the project by using the new technological resources at their own universities even though the time limit set for this project has been finished and the project has been completed. Further related interactive atlas projects will be able to benefit from these experiences and developments. Many students use this opportunity to use the findings in their graduation papers. The parts of this project that need to be supported for development are summarised below:

Design: Presenting information on screen is much more limited than on paper. But with a use of a good database and the ability to present numerous comparative data along with each other is easily possible. It is possible to show the same information in different forms for easier understanding. Maps being the best tool for comparative presentation of data, rules for designing computerised maps must be set. How to use technological development in this field must be searched. The existing software is not adequate even for classic cartographic rules.

Animation: Atlases are the most suitable tools for comparing maps in geographical perception means. But classic atlases cannot fulfil this as needed because of their consecutive formation. Electronic atlases have overcome this problem. With widespread use of animated maps spatial changes against time or other factors can be more understandable in presentation. The development of computer and multimedia programming has enabled animation to be used more freely in cartographic presentations. Programs that enable the presentation of data through animated and interactive maps must be developed. The use of 3DStudioMAX software for cartographic purposes was tried in this study with some samples and other means have been studied.

National Atlas: In order to increase the number of electronic atlases such as this, there is a need for valid topographic base maps and a regular flow of spatial data regarding the concerning elements. For organising the data/information that are collected from different disciplines there must be a national policy for atlas creation. SIS must work in close cooperation with these disciplines and must take the necessary steps for the data to be useable. Because classical atlases are cultural values for many countries, producing electronic atlases should also be an honour and task of a government. This work must also be realised as teamwork.

REFERENCES

Bar, R. and Sieber, R., "Towards High Standard Interactive Atlases: The GIS and Multimedia Cartography Approach." 19th International Cartographic Conference Proceedings, CD Atlas Section, Ottawa (1999).

Frappier, J. and Williams, D., "An Overview of the National Atlas of Canada." 19th International Cartographic Conference Proceedings, CD Atlas Section, Ottawa (1999).

Gorden, M., "Cartography for Census Purposes." Report of the AD-HOC Commission on Population Cartography, International Cartographic Association, Morelia, Mexico (1987).

Lambrecht, C. and Tzschaschel, S., "National Atlas of Federal Republic of Germany." 19th International Cartographic Conference Proceedings, CD Atlas Section, Ottawa (1999).

Ormeling, F., "Atlas information systems." Proceedings of the 17th ICA/ACI International Cartographic Conference ICC, pp.2127–2133, Barcelona (1995).

Ormeling, F., "Atlas Terminology and Atlas Concept." Proceedings of the International Workshoop/Seminar "Modern Cartography for Navigating the Information Highway", pp.5-13, pub. by International Cartographic Association, Bandung, Indonesia (1997).

Peterson, M.P., "Interactive and Animated Cartography." Prentice Hall, New Jersey (1995).

Peterson, M.P., "Between Reality and Abstraction: Non-Temporal Applications of Cartographic Animation." http://maps.unomaha.edu/AnimArt/article.html (2000).
Schneider, B., "Integration of Analytical GIS-functions in Multimedia Atlas Information Systems." 19th International Cartographic Conference Proceedings, CD Atlas Section, (1999).
Siekierska, E.M., "From the Electronic Atlas System to the Electronic Atlas Products (Electronic Atlas of Canada from the Beginning to the End)." Proceedings of the Seminar on Teaching Electronic Atlases, pub. by International Cartographic Association, pub. by International Cartographic Association, Visegrad, Hungary (1993).
Smith, R.M., "The Business of Electronic Atlas Development." Note42. CAGIS 26/3, pp.214, (1999).
Tyner, J., "Introduction to Thematic Cartography." Prentice Hall, New Jersey (1992).
Ulugtekin, N., "Overview of Turkish School Atlases." II. Turkish Map and Technique Assembly, pp.759-764, (printed in Turkish) 1989, Ankara.
Ulugtekin, N. and Bildirici, I.O., "Advanced Mapping Technologies: Electronic Atlases." Second Turkish-German Joint Geodetic Days, pp.641-649, Berlin (1997).
Ulugtekin, N. and Bildirici, I.O., "A New-Lowcost Approach to National Statistical Electronic Atlas." Third Turkish-German Joint Geodetic Days: Towards a Digital Age, Vol:2 pp.579-588, Istanbul (1999).
Ulugtekin, N., Ucar, D., Bildirici, I.O., Ipbuker, C., Ozerman, U., Gokcen, O.F., "Production of the Statistical Atlas of Turkey Using Electronic Atlas Technology for Educational Purposes: A Case Study." Istanbul Technical University Research Project. No:1185. Istanbul (2000).
Statistical Yearbook of Turkey, State Institute of Statistics, Prime Ministry Republic of Turkey. Ankara 1997.
www.die.gov.tr situation in March 2000
www.dpt.gov.tr situation in March 2000