Finnish Journal of the Surveying Sciences, Volume 21 no. 1-2, 2002 pp. 1-18

The Topographic Database as an integral part of the Finnish National Spatial Infrastructure - Analysis of the present situation and some possibilities for the future

ANTTI JAKOBSSON
Helsinki University of Technology,
Institute of Cartography and Geoinformatics
National Land Survey of Finland
PO BOX 84
FIN-00521 Helsinki
Finland
Tel. +358 20541 5177
Fax. +358 20541 5070
Email: antti.jakobsson@nls.fi

Abstract. In a strategy for national mapping, the Ministry of Agriculture and Forestry has set a goal that the Topographic Database (TDB) would become one of the main sources for topographic information in Finland. This study investigates the present situation using two data products where the TDB has been used. The author has earlier presented a model for a database-driven production paradigm with an organisational model, and the study seems to corroborate the idea. In an example the National Land Survey of Finland (NLS) could introduce individual identifiers for each feature instances.

The study will describe two products that use the TDB. The first is a 3D visualisation of reality by Instrumentointi Ltd, and the second is the SLICES land use database, which is a joint venture between several data producers in Finland. Although these cases cannot be used for determining the total quality, they show that the TDB represents reality quite well.

This article will show that the TDB is one of the building blocks in the Finnish National Spatial Data Infrastructure (NSDI). The examples show that it can be used for various purposes. Further enhancements in the conceptual model and data content should still be made. There is a clear need for combining different data sources using i.e. unique identifiers. The Basic Topographic Framework (BTF) model described by the author could be one solution. Organisations should take the multi-producer environment where players act in different roles into consideration. This means that quality management plays a key role in the production process. Other multi-producer environments such as the DIGIROAD data set and the use of municipal data in the TDB should be studied further. DIGIROAD will establish a national road and street database by 2003 covering all vehicle-accessible roads in the country. Co-operation between national agencies and municipalities should be increased.

Key words. Spatial data infrastructure, reference data, data quality, strategy, topographic mapping, SLICES, 3D visualisation

1 Introduction

The Topographic Database (TDB) now covers almost theentire area of Finland. In 2001, the National Land Survey (NLS) celebrated the completion of the TDB after a 10-year projec, which began in 1991 when the data model for the present topographic database was defined. The TDB is a general, topologically matching, comprehensive and accurate geographic database covering Finland. It is being used by the NLS to build several cartographic and data products. The most famous is the basic map 1:20 000 which recently celebrated its 50th anniversary (original production began in 1948). Another product is the topographic map 1:50 000, and several small-scale databases (1:100 000, 1:250 000, 1:500 000, 1:1000 000 and 1:4500 000) are based on this. It is compatible with the NLS's cadastral boundary data that is also presented in the basic map 1:20 000. The elevation model is based on the contour lines of the basic map so it is compatible with the TDB.

The NLS has been a pioneer in promoting the shared use of geographic information in Finland. In the 1980s, the LIS project in Finland already recognized that geographic information is important to society (RAINIO, 1988). After that, the NLS led the effort to define the National Spatial Data Infrastructure (NSDI). The work was completed in 1996 (RAINIO, 1996). Several actors have been active in promoting geographic information. One effort was a description of the core data sets in Finland by the Ministry of the Interior's Advisory Committee for Information Management in Public Administration (JUHTA, 1998). Following from that, the Ministry of Agriculture and Forestry has prepared a strategy for geographic information in administration (2001) and recently a new strategy for national mapping (VERTANEN, VAJAVAARA, 2002).

The main objective in this strategy for national mapping is the establishment of core topographic data as the basis for the production of national geographic information as illustrated in **Figure 1**.

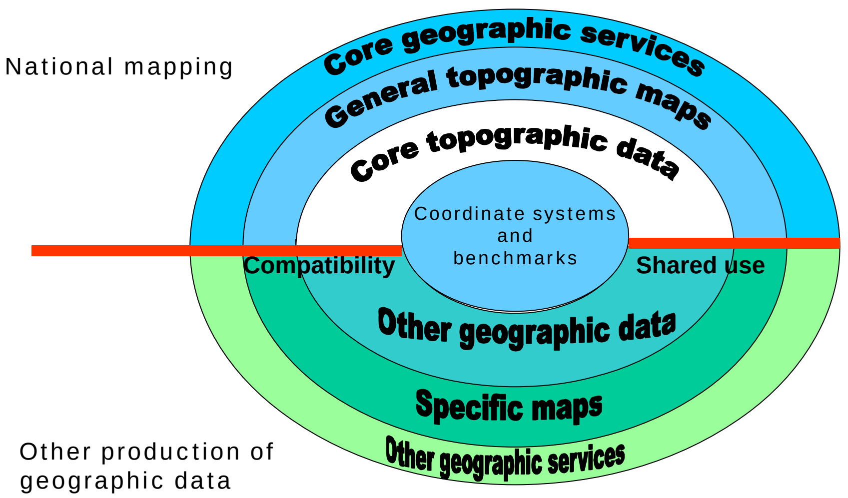

Figure 1 National mapping as a core part of geographic data production (modified from VERTANEN, VAJAVAARA, 2002)

There are several different ways of using the TDB in the production process. These include:

- as part of a printed map product
  - the NLS's own products
    - the basic map 1: 20 000
    - the topographic map 1:50 000
    - small-scale maps (1:500 000)
  - specific maps
    - recreational maps
- as a part of a geographic product
  - raw material with no connection to the original source
    - 1:200 000 road map (Genimap)
    - map databases (1:100 000-1:1:1000 000)

    - as intelligent data with metadata
        - DIGIROAD
        - SLICES
- as part of geographic services
    - backdrop picture
        - the NLS's Mapsite (www.kartta.nls.fi)
        - Genimap.com (a service where you can download topographic maps to a PDA
    - intelligent data
        - DIGIROAD
        - GiMoDig prototype (a research project for mobile use of topographic data sets gimodig.fgi.fi)

The main idea in this study is to explore whether the TDB fulfils the vision of forming the core basic topographic data set for Finland and how it is used especially with some new products and services. The framework for this research is depicted in **Figure 2**. The author has been studying the technology and organisational aspects of the topographic data (JAKOBSSON, SALO-MERTA 2001), introducing a database-driven production paradigm, the basic topographic framework and an organisational model for implementation.

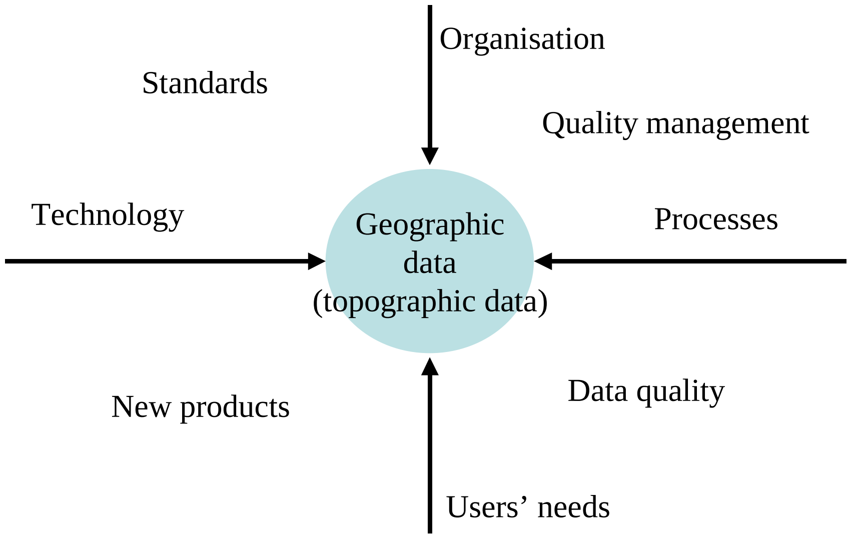

Figure 2. Framework for the research

2 Methods

This study will illustrate some examples of the current use of the Topographic Database with other data sets and analyse the results. There are several relatively new products or possible products that could have been examined in this study. The author selected two examples. The first is a visualisation or a simulation of reality using the TDB, and the other is the SLICES land use data set. Later there is a discussion of other possibilities. The visualisation example was chosen because it offers an excellent opportunity to compare the reality and quality of the data sets. If the TDB is a generic topographic presentation, it should enable a good basis for virtual reality. The SLICES land use data set was chosen because there is a recent study on the accuracy of the data set (HELMINEN, JAAKKOLA, SARJAKOSKI, 2001) and it has some aspects in common with the TDB. It is also a very good example of shared use of geographic information in Finland. Several different data sets were used and it is interesting to study the results.

Using these case studies can give an idea of how generic the TDB is for different applications. Are there any features missing from the TDB at the moment, and is the data quality level adequate for these purposes? The results are compared with earlier studies.

For the visualisation, an area in Valkeala was chosen as an example to demonstrate the capabilities. The NLS has two areas where the simulation has been calculated. Using the simulation software, some screen plots were taken from an area of 200 x 400 m. The area was chosen semi-randomly. The only criterion was that it should contain a typical summer cottage area. The author visited the site and compared the visualisation with the reality using a 1:5000 print out with a typical cartographic presentation of the basic map.

For the second study, a visual comparison of the TDB and the SLICES data was carried out. The SLICES project has examples of the data set on the Internet, and two areas were chosen. Using a colour PDF print-out from the TDB, the two data sets were then compared. After visual comparison the results were confirmed on site. The author visited both sites and compared the datasets against the reality.

3 Case study of an area of Valkeala – visualising the reality

3.1 Content and production

This 3D visualisation is mainly used in military applications. In the future, there could, however, be more uses when computer power and the bandwidth enable this technique to be used on the Internet. A 3D visualisation could be useful for the interpretation of data. It is excellent for visualising possible errors in the field. Using data from the NLS, Instrumentointi (www. instrumentointi .fi) has made an application for simulating reality using the TDB, land classification data and digital elevation model (DEM). It is possible to visualize the whole area of Finland with it. The company has made an application for stationary bicycle in addition to military uses. There a cyclist can select different sceneries derived from the data sets.

3.2 Comparison with reality

An area of 200 x 400 metres was chosen to demonstrate the 3D simulation and compare it with the reality. The area is shown in **Figure 3** with the comparison areas. It is located on map sheet number 311312C3 in Valkeala municipality.

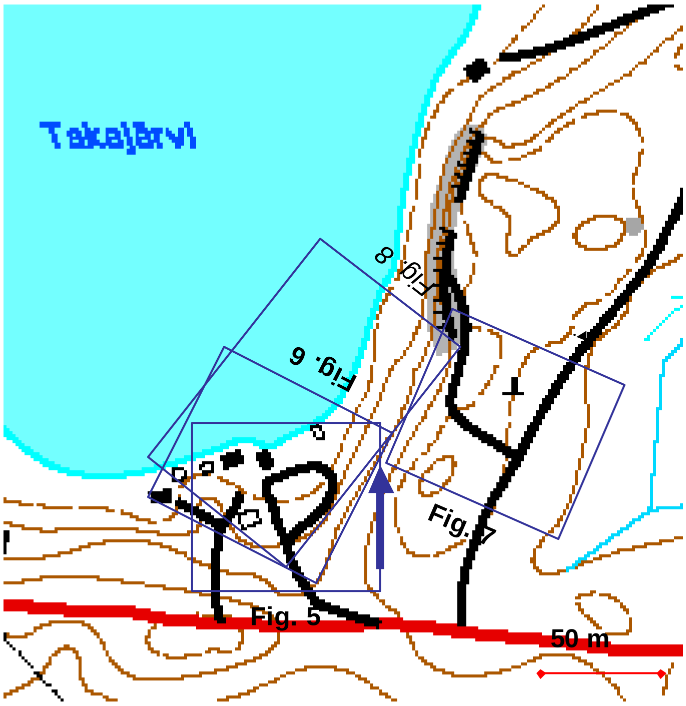

Figure 3. The Topographic Database and comparison points

The same area is then showed in **Figure 4** using a 3D visualisation. Compared with the TDB, one can notice that all the houses, outbuildings and roads are present. Few mistakes were noticed in the field check. According to the TDB, the road that passes the area should be paved, which is incorrect. Also, the road that leads to the first steep (shown on the map) in the forest actually continues to a second one. Otherwise the reality and the TDB match each other.

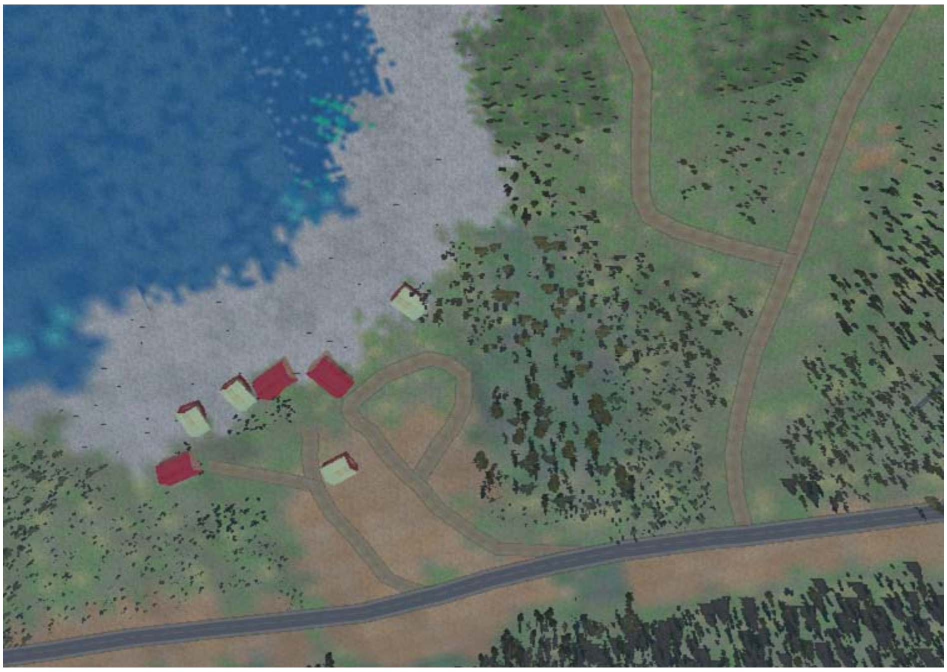

Figure 4. The test area as seen from above

In Figure 5, scenery from the road is compared with the reality. There are no trees in the 3D visualisation. This is probably due to the errors in the land classification data. Also, the terrain looks quite deep towards the lakeside suggesting some errors in the elevation model or in the use of it. Possible sources of error are the pixel sizes used in the elevation model, because the contour lines are correct in the TDB.

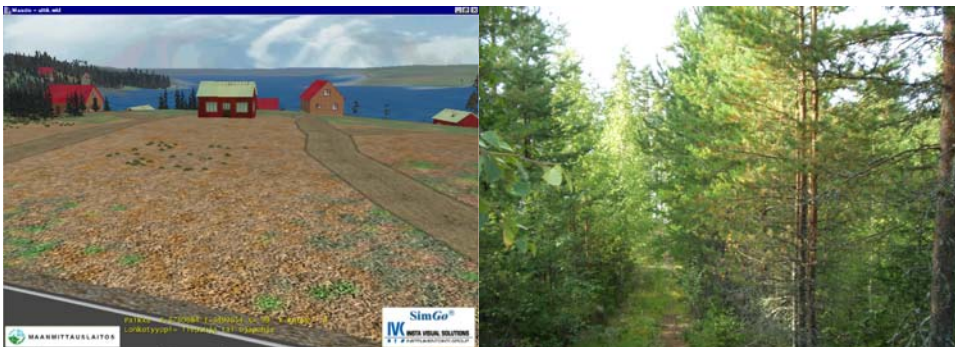

Figure 5. A 3D visualisation compared with the reality

The variation in the elevation is evident in **Figure 6** where there is a summer cottage by the lakeside. It is not clear if it is classified as a summer cottage because the plot of the TDB does not visualise this information. In the simulation, the sizes and the classification of houses are not indicated and a standard texture is used to show a house. The trees in the picture block out the other building, which is quite near, as shown in the visualisation.

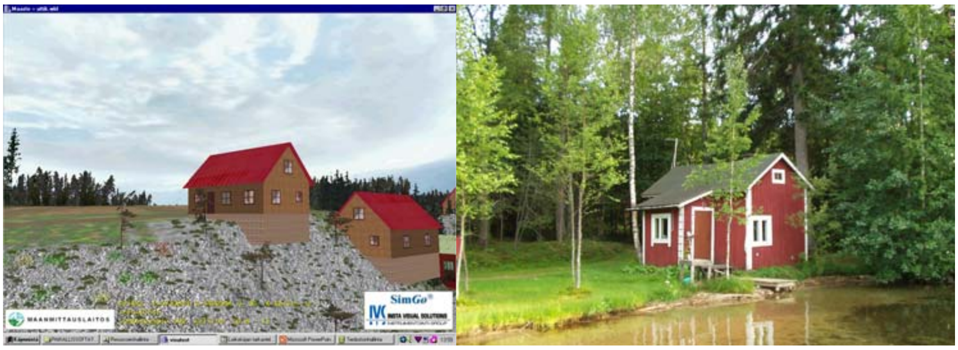

Figure 6. A summer cottage at the lakeside

In **Figure 7** the visualisation of the terrain is very realistic. One can notice that the classification of trees is correct but the extent is wrong. A large stone is missing from the visualisation, but is correctly presented in the TDB.

Figure 7. A forest road

In the last comparison of the reality and 3D visualisation (**Figure 8**) one can notice that it is hard to see other objects when you are in the middle of the forest. Using a visualisation package you can set your altitude very easily. The tree classification here is incorrect.

Figure 8. Scenery in a forest

3.3 Conclusions

Although the case cannot be used to determine the total quality of the visualisation, it shows that the TDB represents reality quite well. It contains the most important feature types that are required for visualisation. Correct forest classification would have improved the results the most. At the moment the TDB does not have forest data as polygons. Also, this example shows some need for the 3D modelling of feature types (i.e. buildings and roads).

The accuracy of the TDB is good enough for this type of application. There is evidence that combining the DEM and the TDB has not been successful in this example. The quality of the DEM needs to be more accurate. Further research is needed to determine if it is a process problem or related to data quality. The forest classification showed many errors in position and thematic accuracy. Clearly there is a need for improvements in that respect. One solution could be the introduction of borderlines of forests into the TDB and the use of remote sensing for classification. The road and building information need to include some information on the elevation in order to better fit the ground. This could easily be incorporated into the data updating process. Some evidence of a problem relating to the classification of buildings was noticed. Buildings are classified depending on their use, which is not actually noticeable in reality.

The example uses predefined textures in buildings to simulate the reality. It could be possible to use real photogrammetric textures taken from aerial photogrammetry, but this would probably only help in some instances. The typical buildings in this area are hidden by the forests so it could only be used in open areas and perhaps for tall buildings. Textures from aerial photogrammetry could be used to represent the forest types.

4 Case study SLICES – land use data set

4.1 Content and production

The SLICES (Separate Land Use/Land Cover Information System) project was initiated in 1997 by the Ministry of Agriculture and Forestry and headed by the NLS. The NLS had produced land classification data using satellite imagery and forest classification data from the Finnish Forest Research Institute even earlier than this. The aim of the SLICES project was to establish a system for producing suitable data on Finnish land use, land cover and soil conditions and to meet user requirements (e.g. monitoring the environment). Compilation of test data for the land use database began in 1998, and production of the land use database was finished by the end of 2000.

The basic idea behind SLICES is to offer separate information on land use in the whole of Finland. The elements are Land use, Land cover, Soil and Restrictions. The Land use element has been produced so far.

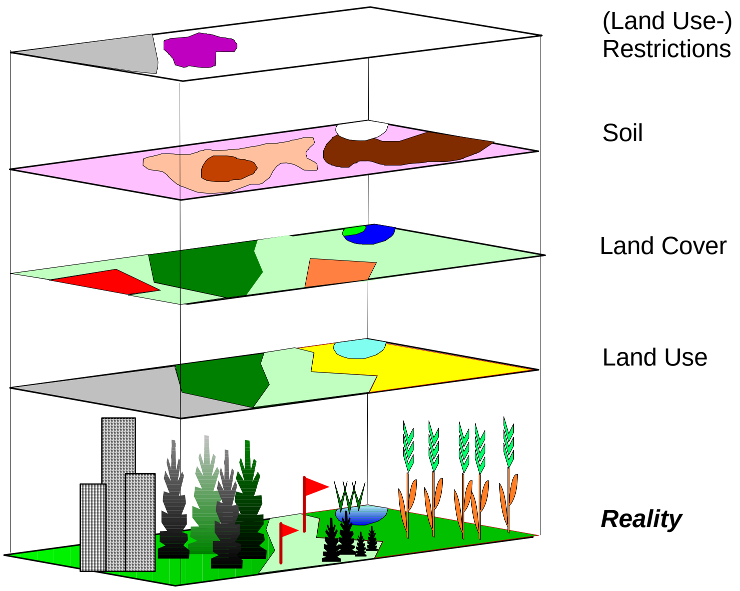

Figure 9 The SLICES data elements (SLICES,2000)

The data sets used for the production of the land use element are described in Table 1.

Table 1 The SLICES classification and the source data sets

Main classification	Source Data/ Producer
A = residential and leisure areas	Building data/ the Finnish Population Register Centre The TDB was used for parks, sports grounds and recreational areas in part of the data set / the National Land Survey of Finland
B = business, administrative, and industrial areas	Building data / the Finnish Population Register Centre
C = supporting activity areas	Roads from the TDB, traffic areas and car parks, building data, other traffic areas / raster basic map and CORINE LC Power lines/FINGRID
D = rock and soil extraction areas	Mainly the TDB, also raster basic map and CORINE LC
E = agricultural land	Mainly the land parcel register, some feature types from the TDB (meadows, fields withdrawn from use (that is missing from the land parcel register)
F = forestry land	Forest classification data/ the Finnish Forest Research Institute
G = other land	Forest classification data/ the Finnish Forest Research Institute and CORINE LC/ the Finnish Environment Institute
H = water areas	Water register, the Finnish Environment Institute

4.2 The visual inspection

Comparing the TDB and the SLICES data visually, one can notice some mistakes in the extent of business, administrative and industrial areas, which are located in the right corner in the Jyväskylä site (**Figure** 10). The extent of the areas seems quite far from the truth. Residential areas seem quite good (both the attribute and the extent information). Some mistakes relating to roads indicate some problems with production, because they are drawn from the TDB. Surprisingly, some meadows seem to be missing. Large areas of water are correct but small areas seem to be either too small or missing.

Figure 10 SLICES and the TDB in Jyväskylä (SLICES, 2000 and the NLS)

In the Pori area there are similar mistakes in the roads. Some fields seem to be missing. Some residential areas seem to be overestimated. Small water areas are missing.

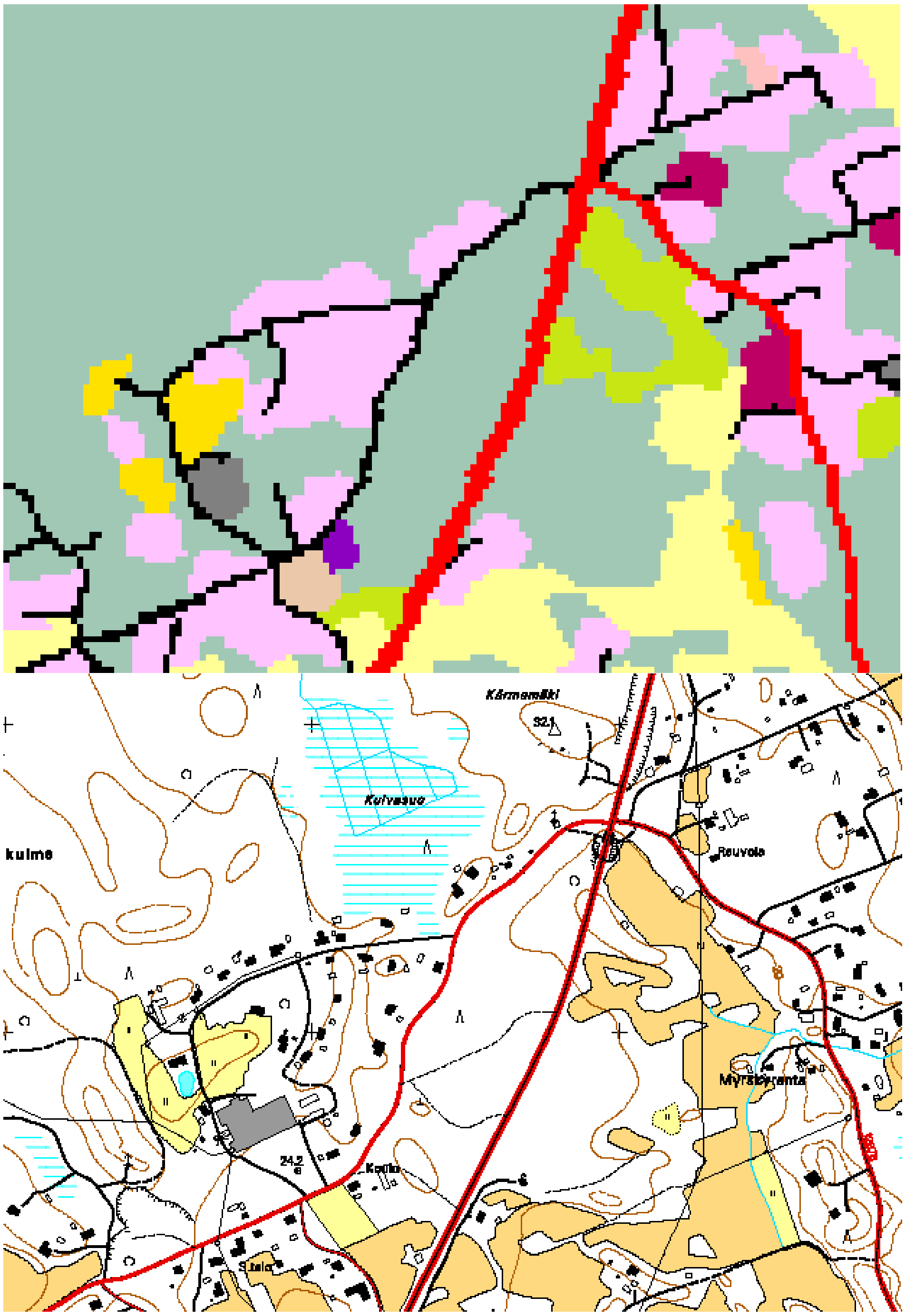

Figure 11 SLICES and the TDB in Pori (SLICES, 2000 and the NLS)

4.3 Field comparison

Both the Jyväskylä and Pori areas were visited by the author in order to confirm the findings with a visual inspection. The field check confirmed the errors in the extent of business, administrative and industrial areas in the Jyväskylä area. This area has very large parking places which should have been classified as supporting areas. An example is depicted in **Figure 12**. In Pori, a small area of water is missing in **Figure** 13.

Figure 12 A parking place

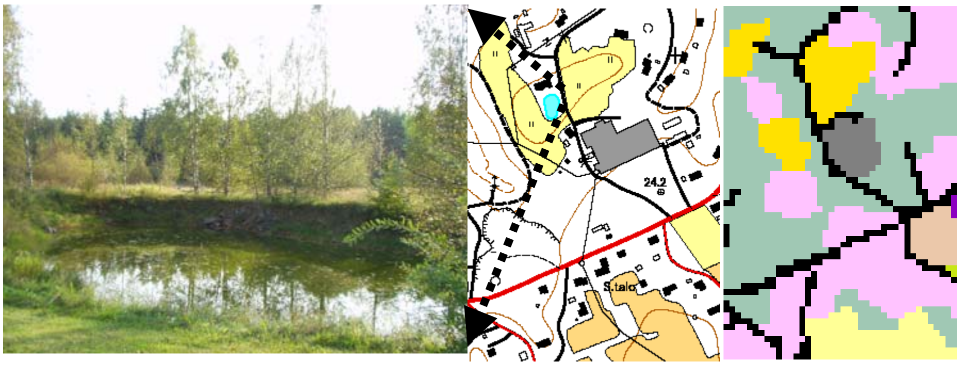

Figure 13 A small water area

Agricultural areas are a good example of where the use of different sources has successful. A land parcel data set has the areas which are in active use, while the information can be outdated in the TDB (as in this case). An example of the fields that are no longer in use is shown in **Figure 14**. A former field is correctly classified in the SLICES data set while it is still classified as a field in the TDB data set.

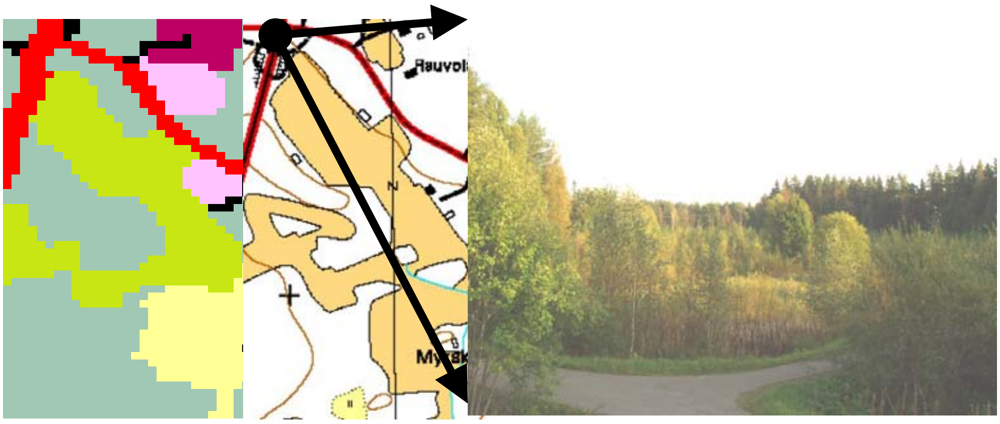

Figure 14 A former field

An example of a slight exaggeration in the SLICES data was the main building area of farms depicted in **Figure 15**.

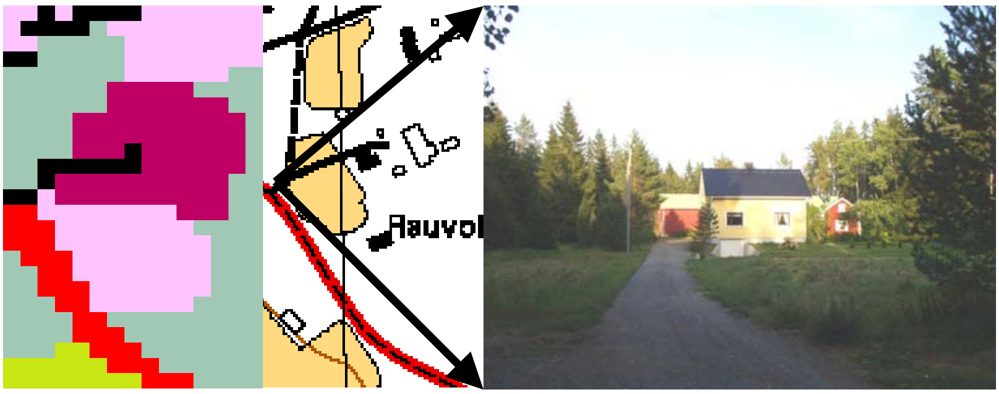

Figure 15 Main building area of a farm

4.4 Conclusions

The SLICES project used the building register as a source for the residential and industrial areas. This is quite a logical choice, because it is the authoritative source for the attribute information on the usage of buildings. On the other hand, the TDB would be quite a logical source for information on the location and extent of buildings. At this moment there is no connection between these two registers, and this example clearly shows a need for a connection. Also, it seems that the land parcel database of the Ministry of Agriculture and Forestry and the TDB agricultural data should be combined or connected with each other. The land parcel register has the most updated information about actively used agricultural areas. There may be a similar need with the water register and the TDB. The water register should at least be updated using the TDB. The examples clearly show that there is a need to combine several different data sources in order to make better data products. One possibility for further investigation would be use of the 1:100 000 small-scale database in the production of the SLICES data. The generalisation level of these products seems to be equal.

The Finnish Geodetic Institute (FGI) has carried out quality control on the SLICES land use data set (HELMINEN, JAAKKOLA, SARJAKOSKI, 2001). They studied 12 different 10 km x 10 km areas and inspected 5810 pixels using stratified random sampling. The reference materials used were orthorectified aerial images, the TDB and field checking. Field checking was used only if there were differences between the interpretations of the TDB and the orthorectified aerial images. They found that 77% of residential areas and 59% of business, administrative and industrial areas were covered in the data set. When residential or business areas were present in the data set, the likelihood that they had been classified correctly was 88% and 90% respectively. These could be reported using two quality elements of ISO 19113: completeness and thematic accuracy. Completeness is defined as presence or absence of features, their attributes and relationships (ISO 19113). In an object data model this would mean that there is an error if a building is missing or if there are too many buildings in the data set compared with the reality. In this case, for example, it would mean main class residential area. The errors in completeness would be omission (23% - 41% features missing) and commission (12% - 10% too many features). The more detailed classes could be reported as thematic accuracy with the misclassification matrix used by the FGI. The use of standard quality indicators would help users to interpret the results in future.

The findings support a connection between the TDB and the building register. In the supporting areas they found that parking places in particular were most likely to be missing from the data set. Again this suggests some needs for improvements in the production process and in the source data sets. In agricultural areas, they found that fields withdrawn from use had the most errors. This again supports findings suggesting the need for a connection between the TDB and the land parcel register. In water areas, they found the omission error was 3% (they call it object accuracy 97%) and commission error 2% (they call it interpretation accuracy 98%). They found small areas of water and rivers were the most problematic.

The visual inspection method used in this study can give an indication of the possible errors. The TDB can be regarded as representing reality quite well. In the quality control of the FGI they visited approximately 15 – 40 pixels in a test area. They do not state the total number inspected at the actual location, but it could be around 6%. This might indicate that the TDB could be used in the future for automatic accuracy assessment. Of course there is still a need for field inspection.

5 Discussion

The conclusions from the SLICES test case suggest that the TDB and the building register, water register and land parcel register should be connected using for example unique identifiers. This has been suggested by the author (JAKOBSSON, SALO-MERTA, 2001) in the Basic Topographic Framework model (BTF). In Great Britain, Ordnance Survey (OS) has started using this approach (OS, 2000 and JUSTHAM,2002) in its MasterMap product.

Similar strategies where a new vision for national mapping have been presented include the US Geological Survey's (USGS) The National Map (2001). One of the goals in their vision is that changes will be captured in near real-time. The ultimate goal is that changes would be recorded within 7 days of a change occurring on the landscape. It also states the data will be available on the Internet. It sets the roles for USGS as being: "the (1) guarantor of national data completeness, consistency, and accuracy, (2) organizer responsible for awareness, availability, and utility of The National Map, (3) catalyst and collaborator for creating and stimulating partnerships, (4) integrator and certifier of data from participants, (5) owner and data producer when no other sources for needed data exist, and (6) leader in the development of geospatial data standards." In the organisational model of the BTF (JAKOBSSON, SALO-MERTA 2001) the author presented that one organisation should take responsibility for data management and data quality while others could act in different roles (producers, users). The National Map seems to corroborate this. In the SLICES project there is evidence of similar roles. The NLS is acting as a producer in the project organisation consortium. It seems that the production of data sets is shifting from a single producer model to a multi-producer or consortium model where responsibilities are shared.

The National Map consists of several themes. These are high resolution digital orthorectified imagery (aerial or satellite), high resolution elevation data, vector feature data for the themes of hydrography, transportation (roads, railroads, waterways), structures, boundaries of governmental units, and administrative boundaries of publicly owned lands, geographic names and land cover information. Interestingly, the National Map is dived into themes, while in the Finnish strategy there is a division into different databases (the TDB, map databases, the DEM, aerial photographs). The difference is that the theme structure picks out the most important real-world phenomena for the society. It could allow more flexibility in the production programmes of the future. In Europe, the European Commission (EC) has been active in trying to define the reference data sets. First the EteMII project (HANCOCK, 2001) and now the INSPIRE project have defined the most important data themes. The INSPIRE (Infrastructure for Spatial Information) project aims to make relevant, harmonized and quality geographic information available for the purpose of formulation, implementation, monitoring and evaluation of Community policy-making. One of its working groups has produced a draft working paper where reference data components are described (RASE ET. AL., 2002). These include: geodetic reference system, units of administration, units of property rights (parcels, buldings), addresses, selected topographic themes (hydrography, transport, height), orthoimagery and geographic names.

In the first case study one can notice the importance of linking all these elements together (e.g. buildings, elevation and land cover). From the perspective of different themes and the BTF model one could feasibly combine all these elements under the same umbrella, as shown in **Figure 16**. Using the BTF, one could easily update the geographic information needed in several registers. A virtual unification of the different data themes and data sets could be performed. On the other hand one could make several products and services using the thematic information and the BTF.

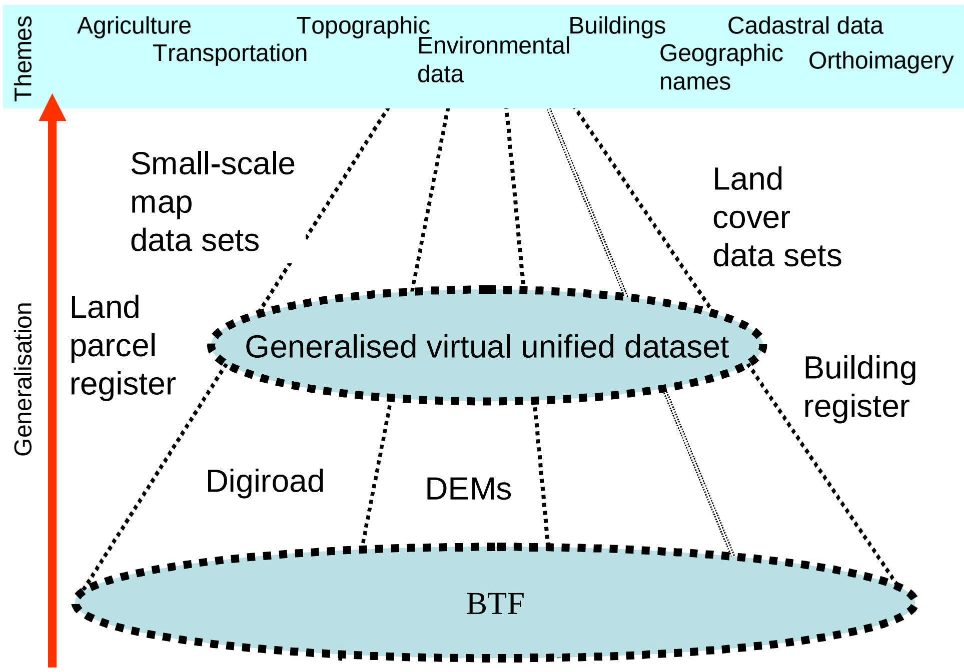

Figure 16. The reference data themes and the Basic Topographic Framework (BTF)

The US version sets more definite goals for how up-to-date the information should be compared with the Finnish National Mapping Strategy. It presents variable resolutions for different areas and themes. This is quite similar to the BTF model discussed earlier.

In 2001, a new advisory board for geographic informationwas set up within the Ministry of Agriculture and Forestr. This new board will study the various options for increasing cooperation between producers and users. It will define the new strategy for the Finnish National Spatial Data Infrastructure. It will be interesting to see if the EC legislation will be established in Europe as a result of INSPIRE proposals.

6 Concluding remarks

The TDB is one building block within the Finnish NSDI. The examples show that it can be used for various purposes. However, some enhancements should be made in the conceptual model and data content. There is a clear need to combine different data sources using, for example, unique identifiers. The BTF model described by the author could be one solution. Organisations should take a multi-producer model where players are acting in different roles into consideration. This means that quality management is a key aspect of the production process. Other multi-producer environments such as the DIGIROAD data set and the use of municipal data in the TDB should also be studied further. By 2003, DIGIROAD will set up a national road and street database covering all vehicle-accessible roads in the country. Cooperation between national agencies and the municipalities should be increased. One interesting concept that should be studied further is the GiMoDig project (SARJAKOSKI ET. AL., 2002) which explores the possibility of real-time integration and the real-time generalisation of topographic data sets in mobile environment.

The examples showed that visual inspection of data quality can be used to identify problem areas. Some ideas for enhancements in the production process of both products were identified. These should be studied more thoroughly.

Acknowledgements. The author is grateful for the information provided by the SLICES consortium, the National Land Survey of Finland and Instrumentointi Ltd for the usage of data sets. Mr. Aaro Mikkola kindly offered assistance in preparing of the figures presenting SLICES data sets.

References

HANCOCK, T. (2001) Refrerence Data White Paper (v.1.1a), ETeMII project consortium, 200, 48 pp.

HELMINEN, V., JAAKKOLA O., SARJAKOSKI T. (2001) Accuracy assessment of SLICES land use database, Research notes of the Finnish Geodetic Institute 23, 57 pp. in Finnish,

http://www.fgi.fi/osastot/karto/projektit/slices/index_eng.html.

JAKOBSSON, A., SALO-MERTA LEENA (2001). Definition of a Basic Topographic Framework for National GI Policy – One Database for All Basic Topographic Data, In Proceedings of the 20th International Cartographic Conference, Volume 4, pp. 2197-2205.

JUHTA (1998). Core Geographic Data Sets in Finland, final report, Published by JUHTA 3/1998, p. 56, in Finnish, http://194.89.205.3/juhta/julkaisut/PTY2_98.htm.

JUSTHAM, J.(2002). MasterMap within your grasp, *In Geo:connexion*, volume 1 issue 5, May 2002, pp. 44-46.

OS (2000). Digital National Framework, An Introduction, Consultation paper 1/2000, Ordnance Survey, March 2000.

OS (2000). Digital Identifiers, Consultation paper 3/2000, Ordnance Survey, April 2000

RAINIO, A. ed. (1996). National Geographic Information Infrastructure of Finland – starting point and future objectives for the information society, http://www.nls.fi/ptk/infrastructure/index.html.

RAINIO A. ed. (1988). The shared use of geographic information in Finland, LIS project, Publication of the National Land Survey no 60, appendixes A-F, Valtion painatuskeskus, Helsinki, 60 pp. in Finnish.

RASE, D., BJÖRNSSON A., PROBERT, M., HAUPT, M-F. EDS. (2002). Reference Data and Metadata Position Paper, Eurostat, http://www.ec-gis.org/inspire/reports/DPP/RDM_Position_Paper_v3-0_en.doc.

SARJAKOSKI, T., SARJAKOSKI, L. T., LEHTO, L., SESTER, M., ILLERT, A., NISSEN, F., RYSTEDT, R. AND R. RUOTSALAINEN (2002). Geospatial Info-mobility Services - A Challenge for National Mapping Agencies. In the *Proceedings of the Commission IV 2002 Symposium*, "GeoSpatial Theory, Processing and Applications", Ottawa, Canada, 8-12 July 2002, Vol. 34, Part 4.

SLICES (2000). SLICES documentation in www.slices.nls.fi

VERTANEN A., VAJAVAARA R. eds. (2002). Maps and Geographic Information – part of a viable society, Strategy for National Mapping 2001-2010, Publications of the Ministry of Agriculture and Forestry 2/2002, Helsinki, 2002, 28 pp.

VERTANEN A., VAJAVAARA R. eds. (2001). The Geographic Information Strategy for the administration within the Ministry of Agriculture and Forestry – follow-up and directions 2001, Publications of the Ministry of Agriculture and Forestry 3/2001, Helsinki, 2002, 25 pp. in Finnish.

U.S. GEOLOGICAL SURVEY (2001). The National Map: Topographic Mapping for the 21st Century: Reston, Va., Office of the Associate Director for Geography, U.S. Geological Survey.