DIGITAL UPDATES AT THE DUTCH TOPOGRAPHIC SERVICE

Paul van Asperen
Research Engineer
Topographic Service
The Netherlands

Commission IV, Working Group 3

KEY WORDS: Updating, Revision, Database, Vector, Digital Mapping, Change-only, Monoplotting

ABSTRACT

The Dutch Topographic Service (TDN) is the national mapping agency of The Netherlands. It produces maps and databases at scales ranging from 1:10,000 to 1:500,000. The database 1:10,000, TOP10vector, is a topologically structured vector database. TOP10vector is based on the analogue-digital conversion of the conventional map 1:10,000. After the conversion, TOP10vector is updated in a digital way. This causes changes in the production method and offers new opportunities for data-delivery to TOP10vector-users. This article reflects the first experiences of TDN regarding the digital revision of TOP10vector. The production method has been successfully modified: change-only plots are, for example, introduced. However, data-delivery through change-only datasets to users is still an issue to be looked into.

1. TOP10VECTOR: PRODUCT PHILOSOPHY

Figure 1 Map 1:10,000

TDN has been producing the map series 1:10,000 since 1952. An example of this map is given in figure 1. The monochrome map forms also the basis for the full colour map 1:25,000. In 1990, it was decided to switch from analogue map production to digital database production. The actual production of the database 1:10,000 started in 1991. This new product is called TOP10vector. TOP10vector is a digital topographic database in vector-format, which can be used in the scale-range 1:5,000 - 1:25,000. Analogue maps at scales 1:10,000 and 1:25,000 are derived from the database.

In the early nineties, it was difficult to set up the specifications for a product like TOP10vector. User-requirements could not be formulated because experienced digital data-users did not exist. Therefore, three main starting points were formulated:

- TOP10vector should be kept simple in order to be useful for as many users as possible
- TOP10vector should be compatible with as many hard- and software-systems as possible, especially GIS
- the maps 1:10,000 and 1:25,000 should be produced directly from TOP10vector.

As a result, TOP10vector contains all map features that appear on the maps 1:10,000 and 1:25,000. To comply with the second starting point, it was decided to structure the vector data in a topological way, which facilitates the use of the data in GIS. It was also decided to produce TOP10vector in a CAD-environment, in this case Microstation-software running on Intergraph-workstations, because the geometrical objects and their attributes can be produced efficiently in this environment. No relational attribute database would be used. This decision needs more clarification, because some other national mapping agencies have decided to produce their databases in a GIS-environment.

The philosophy behind TDN's decision is that all kinds of GIS, containing user-specific objects and attributes, have to be supported. Therefore, data interchange is crucial and not the data production-environment. Secondly, TDN itself does not have the intention at this moment to query or analyse its data. TDN's application is map production, which does not require a GIS.

The disadvantage of production in a GIS-environment is its complexity. Not only topological relationships have to be ensured, also linkages to the attribute data have to be maintained. Another disadvantage of GIS-production concerns the costs: GIS-licenses are more expensive than CAD-licenses. Obviously, the number of attributes to be stored in a CAD-system is limited. To overcome this problem, each feature can be assigned to an unique identifier, which refers to an ASCII-data-table. Moreover, when the need arises, a GIS-environment can be created for certain applications.

Although TOP10vector is produced in a CAD-environment, a few attributes can be stored, mainly to describe the features in the database. TOP10vector contains coded points, lines and areas. Multiple-coding is applied in order to store the necessary attributes. These attributes are stored within the applicable geometric features. The data is structured in a topological way: all areas are closed and left/right information is added as attribute data.

One TOP10vector-dataset covers an area of 62.5 km^2. The term dataset is preferred over database and is defined as the collection of all topographic files covering one map sheet 1:10,000. A dataset consists of the following files:

- BASE-file (BAS):
  coded point and line features, topologically structured (roadsides, ditches), plus centroïds, representing the area code (forest, grassland)
- SYMBOLS-file (SYM):
  point features together with their cartographic representation, like post-office, bridge
- PATTERNING-file (PAT):
  line features together with their cartographic representation, like row of trees, height differences
- BUILDINGS-file (BLD):
  closed polygons representing individual buildings
- AREA-file (AREA):
  closed polygons, computed from the BASE-file.

The AREA-file is fully derived from the BAS-file. Through the centroïds and their bounding lines, areas are built which are stored as closed polygons with the area-codes copied from the centroïds. The buildings in the BLD-file are not topologically linked with the features in the BAS-file.

A simple example of the content of TOP10vector is given in figure 2. The centroïds of the BAS-file are not visualized.

The Netherlands Council for Geographic Information (RAVI) carried out a feasibility-study for a nationwide uniform geometric reference at scale 1:10,000, a so-called core database. The study revealed that there is a real need for such a database. The defined content and structure corresponds with a selection from TOP10vector-features. Recently, TDN and RAVI have established a TOP10vector-user-forum, in which all interested users can participate. In this forum, user-requirements can be formulated and, when accepted, implemented in TOP10vector. Important users are various ministries, provinces, municipalities, water boards and utility companies.

2. TOP10VECTOR: A/D-CONVERSION

TDN has a long tradition in regularly updating topographic maps. In order to keep the map production going, TDN decided to combine the updating with the analogue-digital conversion. The production method is visualized in figure 3.

TOP10vector is based on an intensive reconnaissance survey, using monochrome aerial photographs at scale 1:18,000. These aerial photographs are enlarged to scale 1:10,000. The process starts with a comparison between the existing map and the aerial photograph. A positive film, representing the existing map, is made. Map features, which are not longer existing, are marked on the film. New topography is annotated on the photograph. Not all changed topography can be interpreted. Also additional information, like land cover classification and geographic names, has to be collected. Therefore, topographers carry out a reconnaissance survey after the preparation in the office. In the field, the same procedure is followed: non-existing map features are marked on the

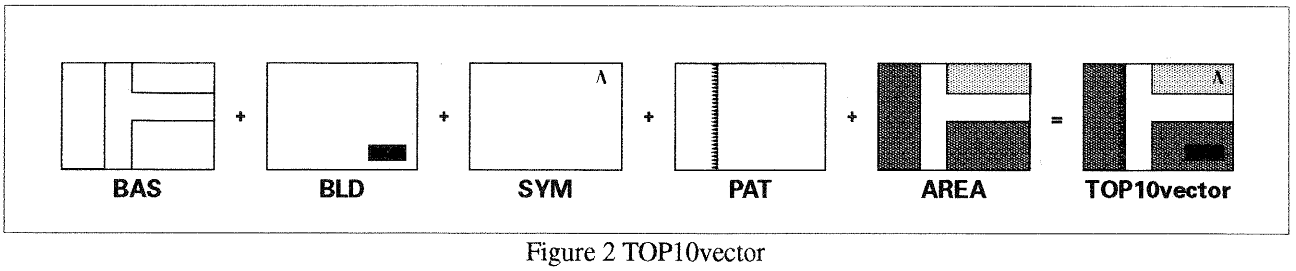

Figure 2 TOP10vector

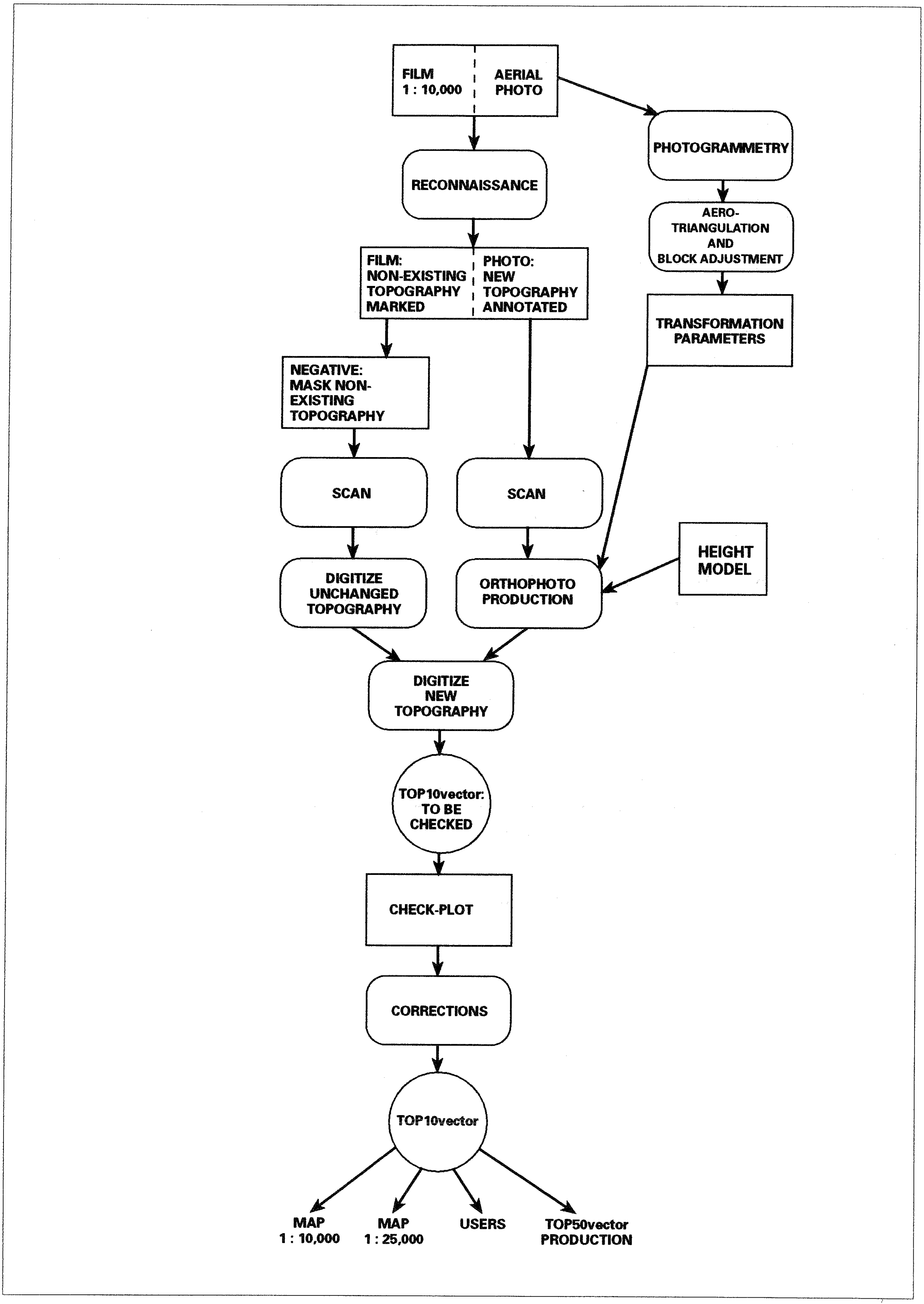

Figure 3 TOP10vector: analogue-digital conversion

film, new topography is annotated on the photograph.

While the topographers are collecting field changes, photogrammetry is applied to provide a solid geometric base for the digitizing process. The original photographs are subject to aerial triangulation and block adjustment. Ridges of roofs are used to establish the relation between ground coordinates and photo-coordinates. The fiducial marks are also measured in photo-coordinates in order to determine their imaginary ground coordinates.

When the topographers return from fieldwork, the enlarged aerial photographs are scanned with 1 meter resolution. The pixel-coordinates of the fiducial marks are measured. Using the ground coordinates of these points, determined in the photogrammetric process, the transformation parameters for rectification can be determined. The control points of the triangulation process are not marked on the enlarged photograph. Therefore the fiducial marks are used. Combining the transformation parameters with a height model (25 m grid), the scanned photograph is transformed into a digital orthophoto. The digital orthophoto contains the annotations made by the topographer.

Meanwhile, a negative of the existing map is made. The marked features on the positive are masked on the negative. This negative then contains the unchanged topography. The negative is scanned and digitized in vector format. This work is contracted out. As a result, a vector-file containing unchanged topography is produced.

Heads-up digitizing is applied to add the new topography to the vector-file. The digital orthophoto serves as a background image. The operator digitizes the annotations from the orthophoto. The operator does not need to interpret the orthophoto itself, because all features to be digitized are annotated with special symbols. These symbols represent the codes which should be attached to the points, lines and centroïds. When digitizing is finished, software is applied to check the data-structure and to built the areas as closed polygons. TOP10vector is then finished, although it needs to be checked. A check-plot is made, which is compared with the markings on the film and the annotations on the photograph in order to check whether all features are correctly digitized. When the corrections are completed, TOP10vector-production is finished.

The TOP10vector-dataset is used for map production at scales 1:10,000 and 1:25,000, and for the production of TOP50vector. TOP50vector is the digital topographic database at scale 1:50,000 and is produced through the interactive generalization of TOP10vector. TOP10vector is also made available to the user community.

At the beginning of 1996, 60% of The Netherlands is digitally covered in the TOP10vector-structure. The first edition of TOP10vector will be completed by the end of 1997. Then, all 650 1:10,000 map sheets will be available in digital format. All these datasets have to be updated regularly.

3. REVISION: GENERAL ISSUES

When an updating system has to be designed, one should consider the following:

1. revision cycle: continuous or cyclic
2. revision cycle depending on area or object type
3. production method: data capture and digitizing
4. type of update delivery.

3.1. Revision cycle

One can decide to have a continuous update or cyclic update. Most cadastral maps and large scale base maps are continuously updated, often by means of land surveying methods. The date of capture or digitizing is often input. Topographic maps of smaller scales are often revised through aerial photography, which make them more suitable for cyclic revision. However, as large scale databases are being built, the digital format makes it possible to apply the large scale data for continuously updating smaller scale databases.

3.2. Areas or objects

In case of cyclic revision, one can choose to vary the cycle for densely populated areas, where more changes occur, and for sparsely populated areas. It is also possible to update some feature types more often than others, for example to update buildings and roads more often than other topography.

3.3. Production method

When the same system of data collection is used, there is no need to change the production method. However, digital updating offers new opportunities, because it is relatively easy to incorporate digital data from other sources, e.g. changes from larger scale databases. It is also possible to store historic data, more or less invisible in the dataset. An example is the updating system of the Ordnance Survey of Northern Ireland (Gray, 1995).

3.4. Type of update delivery

Updated datasets can be delivered to users. However, it is also possible to deliver change-only datasets to users. The last option is promising for GIS-users, who have collected additional user-specific attribute data. When change-only datasets are delivered, there is no need to enter attribute data for the unchanged features for the second time. Such update delivery requires an unchanged geometry at the user's side, or a system based on unique identifiers, maintained at both the producer's and the user's side.

Change-only information can be produced through graphical comparison of features or through a selection on database entry date. Research has been carried out on update deliveries of large scale databases (Gray, 1995) and some methods are already operational. The Dutch Cadastre has an automatic update delivery system, where users can choose their own update frequency. Change-only information is extracted from the master database through selection on database entry date (Lemmen, van Oosterom, 1995).

4. TOP10VECTOR REVISION

Nowadays, three revision cycli of respectively 4, 6 or 8 years exist for 1:10,000 mapping, depending on the characteristics of the area. From 1997 onwards, a four year revision cycle will be carried out: 25% of The Netherlands will be revised completely every year. As a result, cyclic revision will take place, where all feature types will be updated. TDN does not have the intention to store historic data in the datasets.

This year, TDN started to update the first eight 1:10,000 map sheets on the basis of the existing TOP10vector-datasets. The revision process is subdivided in the following steps:

1. data-collection
2. processing
3. checking
4. delivery

The process is visualized in figure 4.

4.1. Data-collection

Data collection for TOP10vector-revision does not differ from data-collection for the initial TOP10vector-production. A positive film is produced from the existing dataset. Non-existing topography is marked on the film, new topography is annotated on the enlarged aerial photograph. This is done during the preparation stage in the office and during the reconnaissance survey in the field.

4.2. Processing

Digital orthophotos are produced from the enlarged annotated photographs. The method is already described in paragraph 2. A copy of the existing dataset is used for the update process. The marked features on the film are deleted from the dataset and new features from the orthophoto are added through heads-up digitizing. The date of deletion or addition of a feature is not captured. The year of aerial photography is stored for each feature, but this attribute value is not used to produce change-only datasets. All features, unchanged and new, are checked on validity. Therefore, the whole dataset gets one reference date.

4.3. Checking

During this step, a check is carried out whether all non-existing features are deleted and new features added. Change-only datasets are produced for that purpose. The existing TOP10vector-dataset is automatically compared with its updated version. Change-only information is extracted in order to produce two types of check-plots. One type shows all deleted features, the other all new features. Checking is done by comparing the marked features on the film with the plot showing the deletions and the annotations on the photograph with the plot showing the additions. On the basis of these comparisons, corrections are made, after which TOP10vector-revision is finalized.

4.4. Delivery

The updated dataset can be delivered to users. It is also possible to compare the updated dataset with the old dataset in order to produce change-only datasets for the revision of TOP50vector and for users.

5. CHANGE-ONLY DATASETS

Change-only datasets are produced for check-plots, for delivery of updates to users and for updating TOP50vector. This paragraph describes the method of change-only extraction.

The individual files (BAS, AREA, etc.) of the old and updated datasets are, one by one, automatically compared: unchanged features in terms of geometry and attribute values are deleted in both files, which results in files with deleted and added features. During the comparison, a tolerance value can be used in order to treat features with small geometrical changes as unchanged. Five different reasons can be distinguished why features appear in the change-only datasets:

1. changes in data collection rules and content specifications
2. changes in topography
3. corrections/improvements in the existing dataset
4. changes due to edge-matching of sheets in different years
5. changes due to structural changes (splitting or merging of features).

Changes in rules: It is clear that every change in the content specification will have its effect on the change-only datasets. Changes in specifications should be avoided as much as possible. However, we deal with a new

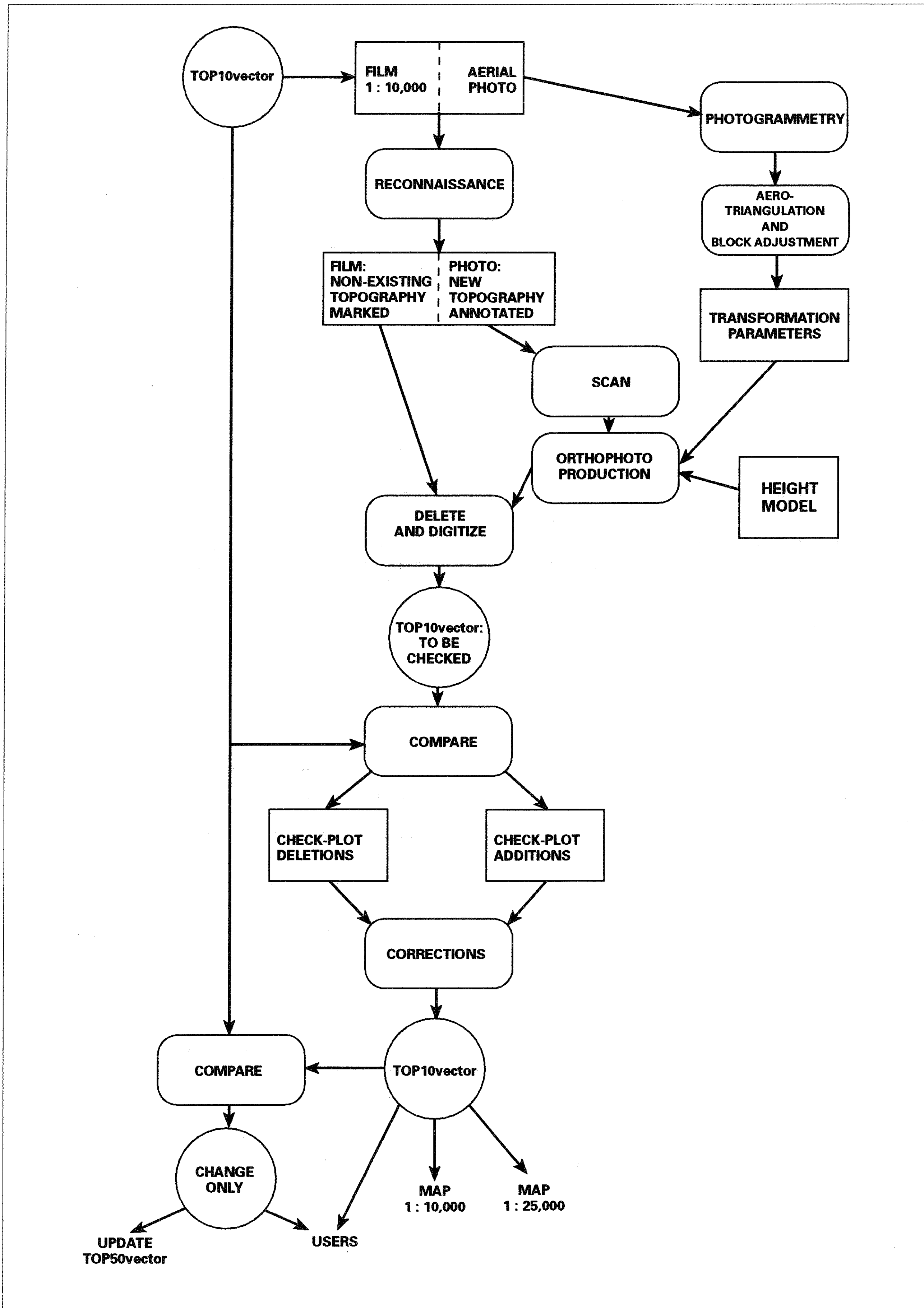

Figure 4 TOP10vector: revision

product for new applications. It is quite likely that specifications have to be adapted to new user-requirements. These requirements may, for example, be formulated within the TOP10vector-user-forum.

Changes in topography: It is evident that these changes appear in the change-only datasets.

Corrections/improvements: This category is of great concern. When the new orthophoto is placed behind the existing dataset, mismatches may be visible. These mismatches may be errors, which have to be corrected. They also may be irrelevant, as they are caused by the statistical processes of the data capture. In our practice, it is very difficult to give clear guidelines to distinguish between errors and statistical noise.

Edge-matching: Edge-matching is the process of connecting features at the border of two adjacent datasets. Connecting datasets produced in different years will result in changes. In the near future, the number of changes caused by edge-matching will drop significantly, because edge-matching has its largest effect at the initial dataset production. However, the problem of edge-matching will always arise at the border of areas with different revision dates.

Structural changes: Due to the topological structure, features will be split up or merged when a connecting feature is added or deleted respectively. The structural changes (split or merge) will appear in the change-only datasets, although they do not represent changes in the field.

It is impossible to produce change-only-datasets where the changes are classified by the five causes of change, described above. Only the structural changes can be automatically filtered. Another valuable difference can be made between geometrical changes and attribute value changes. Therefore, the change-only extraction distinguishes the following types of changes:

a. features-deletions: change in geometry
b. features-additions: change in geometry
c. features-deletions: change in attribute values
d. features-additions: change in attribute values
e. features-deletions without structural changes (subset of a.)
f. features-additions without structural changes (subset of b.)

As said earlier, all file-types are compared. Some types of change are not applicable for certain file-types, due to the content and structure of these files. Table 5 shows which files are produced after comparison of two TOP10vector-versions.

The required content of change-only datasets depends on the application. For the TOP10vector change-only check-plots, the following specifications have been formulated: All changes have to appear on the plots except the following:

1. Structural changes (they are logic and do not need to be checked): both BAS-files FILTERED appear on the plot instead of the BAS-files GEOMETRY.
2. Small geometrical differences, caused by line-smoothing algorithms, should not be plotted. Therefore, a tolerance value of 40 cm is used.
3. Deleted areas (they do not need to be checked): the AREA-files GEOMETRY(del) and ATTRIBUTE(del) do not appear on the plot.

	GEOMETRY		ATTRIBUTE		FILTERED	
	del	add	del	add	del	add
BAS	*	*	*	*	*	*
AREA	*	*	*	*		
BLD	*	*				
SYM	*	*				
PAT	*	*				

Table 5 Files produced by comparison

However, for delivery of change-only datasets to users other specifications may be valid. In order to preserve consistency in the topological structure, they may require all files, except the FILTERED-files, produced without a tolerance value.

Some differences between change-only information for check-plots and delivery to users are visualized in figure 6. The nodes are visualized in TOP10vector to indicate the start- and end-points of line-features. Topographic edges are line features to separate two topographic areas, for example roadsides. In this example four changes are detected:

- a new building
- an attribute-change of an area (cropland into grassland)
- an attribute-change of a line (topographic edge into ditch)
- one area is split into two (grassland split up in grassland and cropland).

Each change will be discussed in more detail, reference is made to table 5 and figure 6.

The new building appears in the BLD-file GEOMETRY(add). It appears both on the checkplot B2 and the change-only file C2 for delivery.

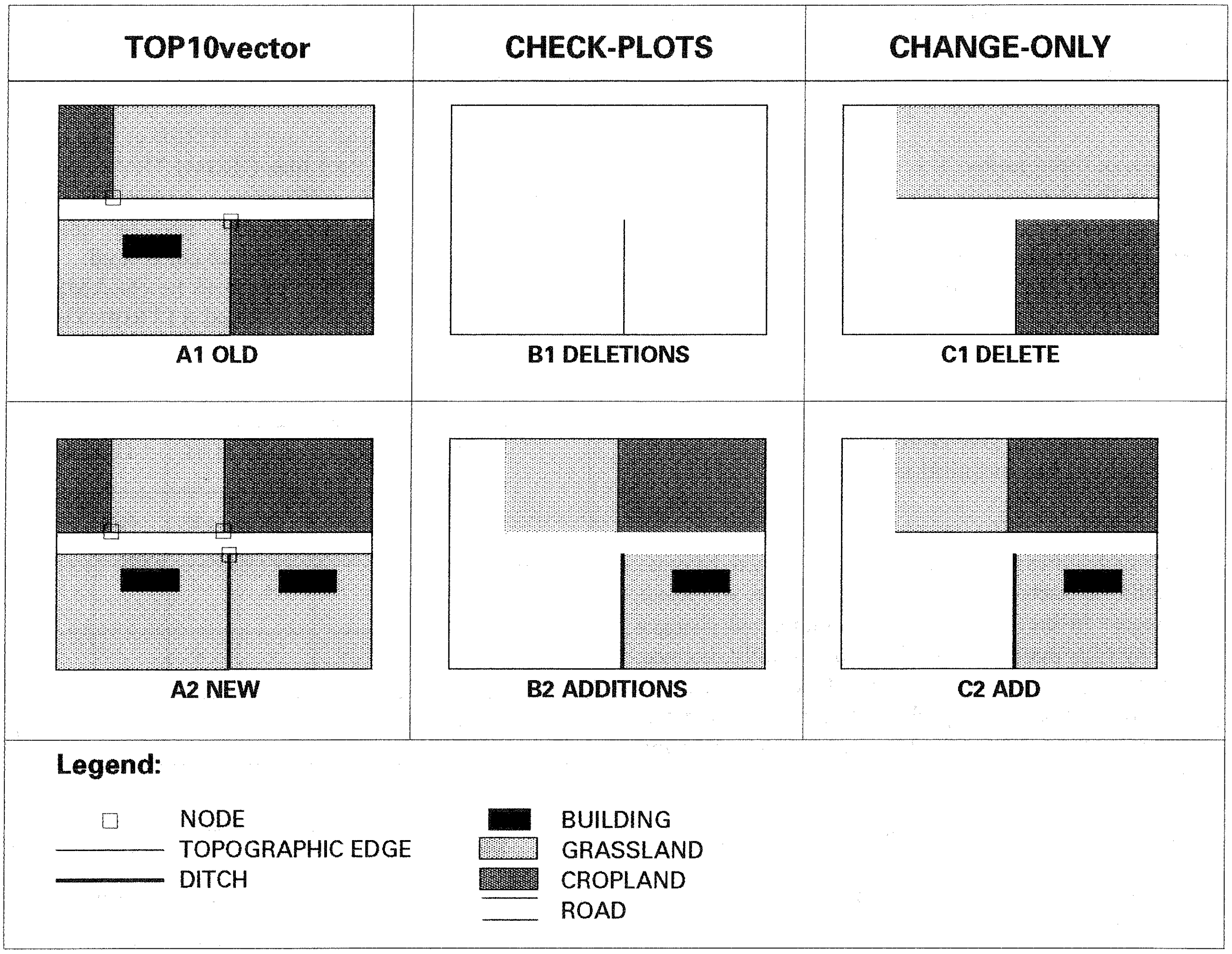

Figure 6 Change-only information

The area on which the new house appears is changed from cropland to grassland. This change appears in both AREA-files ATTRIBUTE. Only the new attribute value (grassland) needs to be checked. Therefore, this area only appears on the checkplot B2. The area appears in both change-only files.

The attribute change of the line appears in both BAS-files ATTRIBUTE and will appear on both checkplots and change-only files. The feature with the old attribute value is plotted, because the topographer has marked this feature on the film during the reconnaissance survey.

The area-split results in changes that appear in both BAS-files GEOMETRY. The structural changes can be filtered to get the BAS-file FILTERED(add), which contains the added edge to split the area, displayed on the checkplot B2. The roadside is split into two edges and do not appear on the checkplots, but are included in the change-only files.

The area-split also results in changes that appear in both AREA-files GEOMETRY. The deleted areas are not plotted. All changed areas are included in the change-only files.

As the example shows, changes in the geometry of line features will cause, of course, changes in areas as well. It is also clear that change-only datasets do not only contain changes caused by real world field changes. The existence of other categories can not be avoided. The TOP10vector-user-forum will be the obvious group to discuss items like this.

In order to facilitate the exchange of change-only information, the application of unique feature identifiers are currently under study. Research and interaction with users is necessary to develop a revision infra-structure which is beneficial to all parties involved.

From a technical point of view, the update process and the production of change-only datasets are rather straightforward. The challenge is to manage and control the process. During the reconnaissance, all topographers should apply the same data collection rules. During digitizing, only changes should be captured. However, it should be permitted to correct mistakes in the datasets. This is a rather delicate issue, because no clear guidelines can be given. Only experience and regularly monitoring can provide tools for efficient update management.

6. EXPERIENCES

One TOP10vector-dataset, based on aerial photography from 1995, has been compared with a previous edition, based on aerial photography from 1991. All files have been compared with a tolerance value of 40 cm. The number of changed features is given in figure 7. Change1 contains changes in the GEOMETRY-files, change2 contains changes in the ATTRIBUTE-files. The total number of changes of the dataset is also given.

FILTERED-changes are not displayed in figure 7. The number of changed features for the FILTERED-deletions is reduced by 20% with reference to GEOMETRY-deletions and for FILTERED-additions, the reduction amounts to 30% with reference to GEOMETRY-additions. The reductions mean a significant acceleration of the checking process.

Figure 7 shows that the total number of features has slightly decreased. Most updated TOP10vector-files have been increased in size, only the number of symbols has dropped enormously. This is caused by a change in the rules: it was decided to omit individual trees. Almost 6,400 individual trees were deleted from the SYM-file. Rows of trees are still captured, they do appear in the PAT-file.

A tolerance value of 40 cm was used, a value which may be reviewed. In case of change-only delivery, a tolerance can not be used as all changes have to be delivered. In this case, when no tolerance value was applied, the numbers of changed features in the BAS-files were increased by 100, which is only a fraction of the total number. Figure 7 then gives an impression of the size of change-only datasets for delivery to users.

Looking at the number of changes, one might question the value of delivering change-only data-sets to users. It is expected that the average number of changes will drop. Three considerations support this expectation:

1. The dataset was one of the first TOP10vector-datasets and contained therefore more inconsistencies than average.
2. During the development of TOP10vector the specifications were altered significantly.
3. The number of field changes in the area is more than average.

Generally spoken, about 45% of the dataset has been changed. The digitizing time for updating is estimated at 50% with reference to the digitizing time of the initial TOP10vector-production. It is expected that this percentage will drop when more experience is gained and when the update process is fine-tuned. The reduced workload due to updating is compensated by the higher update frequency (from 4, 6 or 8 years, depending on the area to 4 years overall).

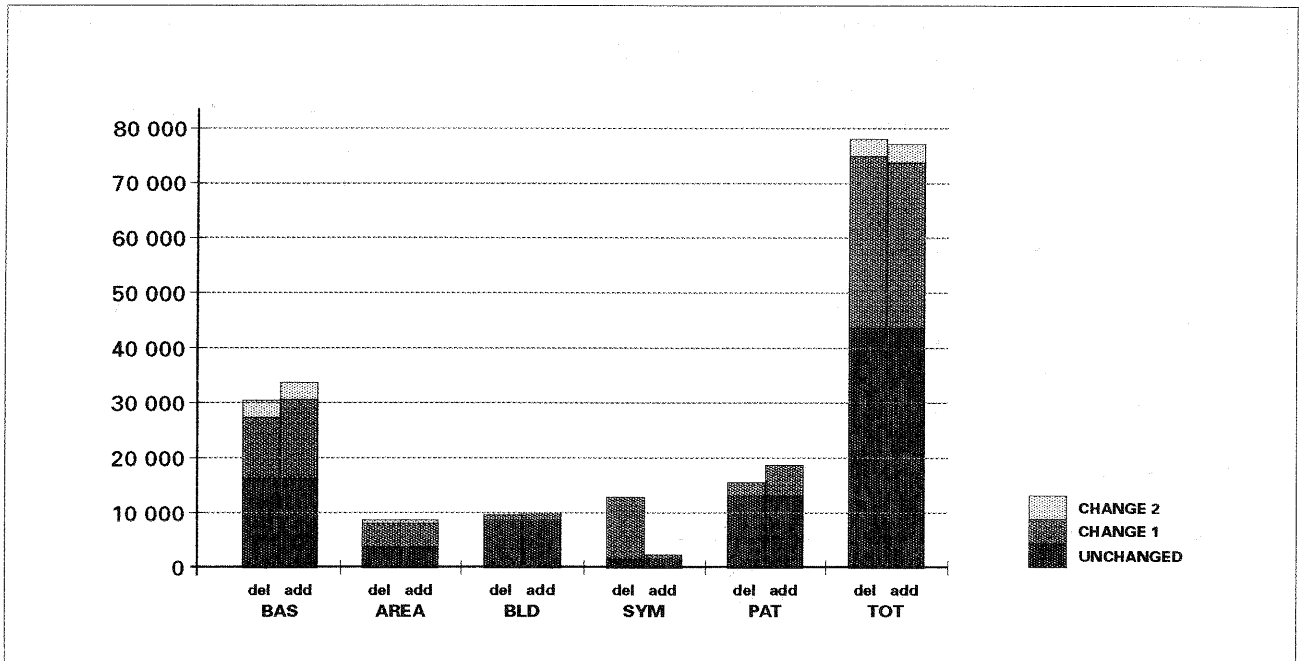

Figure 7 Number of feature changes

7. ALTERNATIVE METHODS

Although the described system is successfully implemented, alternative methods are continuously considered to achieve an efficient production system. Alternatives can be divided into alternative data sources and new technologies.

7.1 Updating through alternative sources

Field changes are captured using aerial photographs. However, most topographic changes are captured by other institutes as well. In The Netherlands, a Large Scale Base Map at scale 1:1,000 exists, which is maintained by regional public-private partnerships. This map is continuously updated and may serve as input for TOP10vector. TDN is investigating this method in order to shorten the revision cycle.

7.2 Updating through new technologies

Other methods of updating are also under consideration. The first one is updating using stereo-workstations. The topographers could use these workstations in order to capture field changes behind their desk and carry out a short field completion. This has an enormous impact on the organisation. Advantages and disadvantages have to be considered seriously.

Another method is updating in the field using notebook-computers. Orthophotos can be loaded and the topographer can digitize the field changes directly in the field. This will be investigated in the near future. The disadvantage of this method is the high demand on photogrammetry during a short period of time. Before the topographers can go for the reconnaissance survey in the summer time, all aerial photographs (covering 25% of the country) should have been transferred into digital orthophotos.

8. CONCLUSION

According to TDN, digital updating in itself is not considered very different compared to the initial TOP10vector-production. The challenge is to control the process. Topographers and operators should work in an uniform way. Another challenge is to deliver change-only datasets to users in order to facilitate their data revision.

At the same time, users ask already for higher update frequencies. Their wishes can be fulfilled by using data from other sources to update TOP10vector. However, for quality assurance, the planned four year revision cycle based on aerial photography will continue to exist.

At the end of 1997, TDN will reach a milestone by having a nationwide digital coverage at scale 1:10,000. Digital updating will then be the normal production method, while new technologies will lead to other innovations of TOP10vector-production.

9. ACKNOWLEDGEMENT

The professional advice and support of Bert Kolk is greatly appreciated. I also would like to thank Jeroen van Gelder for the illustrations.

10. REFERENCES

Van Asperen, P.C.M., 1996, Topographic Databases in Europe - National Mapping Agencies Go Their Own Way, Geomatics Info Magazine, Vol. 10, nr. 2, pp. 17-19

Birth, K., 1995, ATKIS database revision in Nordrhein-Westfalen, In: Workshop on revision of digital topographic databases, CERCO Working Group IX, Emmen, The Netherlands

Geudeke, P.W., 1993, From base database to core database - the Topographic Service as producer of the core database 1:10,000, Kartografisch Tijdschrift, Vol. 19, nr. 2, pp. 24-28 (in Dutch)

Gray, S. (ed.), 1995, Updating of complex digital topographic databases, OEEPE, Official Publication No. 30

Kolk, E., 1995, On the road with TOP10vector - a nationwide medium scale geographic database, NGT Geodesia, Vol. 37, nr. 1, pp. 13-19 (in Dutch)

Lemmen, H.J., van Oosterom, P.J.M., 1995, Efficient and Automatic Production of Periodic Updates of Cadastral Maps, In: Joint European Conference and Exhibition on Geographical Information, The Hague, Vol. 2, pp. 137-142

Opie, R., 1995, The Dutch Topographic Survey: form ministry to marketplace, GIS Europe, Vol. 4, nr. 3, pp. 17-19