# UPDATING AND SPECIFICATION CHANGES - STATISTICAL ESTIMATION OF WORKLOAD

Dorthe Holme and Peter Højholt

National Survey and Cadastre
Rentemestevej 8
2400 København NV
dh@kms.dk and phs@kms.dk

**Commission IV, IC WG II/IV**

**KEY WORDS**: Top10dk, Updating methods, workload, specification changes

**ABSTRACT:**

The vector dataset Top10dk is updated incrementally and new features are added as single objects and objects that are no longer valid are made historic. Single objects can be revised by choosing one of two methods depending on what is considered most suitable in the specific situation.

This paper will look into the usage of the two methods and conclude how the two methods are applied in the production of revision data for Top10dk. The paper analyzes statistical material based on the last three years production of revisions. Based on these statistics estimates of the extend of future updates are made.

Changes within the specification of Top10dk cause a changed workload within the production of revision data. The changed workload is assessed as a correlation between a specification change and the number of revision objects needed to implement the specification change. The experienced gained from the present investigation suggests that a statistical production monitoring system would benefit future workload estimates.

## 1. INTRODUCTION

Top10dk is the national topographic vector-database covering the entire Denmark in scale 1:10000. All data produced and updated adhere to strict topological rules. The initial establishment of Top10dk was finalized December 2001.

Top10dk can be separated in three main parts: the topographic base data, the place names and points of interest and the digital terrain model. This article will cover only the part related to the topographic base data.

## 2. SPECIFICATION CHANGES

The specification of Top10dk has changed over the years from 1995 when the first edition was written and up till 1999 when the sixth edition was published. The first specification had 40 feature types and the last one 51 feature types to map.

The specification has changed to accommodate comments and needs from both the users and private producers of Top10dk. Besides adding new features one of the main improvements of the specification for Top10dk has been the topology table, where the relationship between the different features is established.

## 3. UPDATING OF TOP10DK

The first updating on Top10dk was made based on aerial photos from 1998 in the region of Frederiksborg. From 1999 the full-scale production of updates for Top10dk was started. The area in an annual production covers approximately 20 percent of the country. In September 2001 42 % of the country have been updated.

The production of updates is organized as EU tenders and an annual production is typically split into 5 different contracts. Each contract is divided into two or three sub-areas with an app. size of 80000 ha.

The contract holder supplies KMS with data as photogrammetric models for control and data later data is sampled into 10 x 10 km blocks.

## 4. METHODOLOGY USED IN UPDATING

When updating Top10dk there are two basic methodologies to choose between. The first method is the object method where an object is either completely deleted or added. Changes are registered as a deletion of one object followed by an insertion of a new object. The object method can be applied on all feature types irrespective whether it is a point, line or polygon object.

The partial method can delete, change or add only part of an object. The method can only be applied to features of line or polygon types.

## 5. STATISTICS FOR THE USE OF THE PARTIAL REVISION METHOD COMPARED TO THE OBJECT REVISION METHOD.

5 different 10 x10 km blocks were selected for closer statistical investigation. The selected areas covered a range of specification changes and data from different data producers. The areas had been updated at varying intervals from 3 to 5 years. In the statistical analysis this was taken into account by representing the figures as changes pr. year.

It was found that in all areas the operators had used both the object method and the partial method for all types of revisions. Experienced operators across different companies seem to make similar choices of method.

But there were significant differences depending on object feature type.

For revisions of buildings both method was used to an almost equal extent. It is assumed that the low number of points in each building makes the methods evenly applicable.

For built-up areas operators almost always selected revision by the partial method. In these objects only small changes in few places along a (long) perimeter makes this method the fastest.

For small lakes and other similar labile boundaries operators most often chose the object method, reflecting the fact that almost the entire perimeter of the objects had to be redone.

## 6. CHANGES IN WORK LOAD CAUSED BY SPECIFICATION CHANGES.

Three different kinds of specification changes are identified:

- new features
- changes in the registration of existing features
- recoding part of a feature class to a new feature class

The new feature class that caused the most additional revision data was road boundaries. It increased the amount of revision data with approximately 15-20 %, and the amount of objects in the database by 10%.

The most significant example of changes in the registration of existing features was changes made in the registration of built-up areas. This change was present in two of the selected areas. The interpretation of statistical data for these cases will be included in the final paper.

Recoding parts of a feature class to a new feature class has caused no significant changes in the amount of revision data.

## 7. STATISTICS FOR CHANGES NOT RELATED TO CHANGES IN THE SPECIFICATIONS

These changes fall into two groups:

changes caused by real world changes
changes caused by errors in the initial registration.

Because of improvements in the error detection tools and routines, the level of errors detected in the production process has increased dramatically during the 6 years period covered by the selected data,.

At present it has not been possible to separate data for the two kinds of changes. Depending on the object feature change levels varies from 1% to 15 % pr. year, and it is at the moment difficult to reach a general conclusion.

## 8. CONCLUSIONS

Both the object method and the partial method of updating are used and it is justified to maintain both as parallel choices.

Including new major features in the dataset can increase the amount of revised objects substantially while most minor new features have almost no significance.

To estimate the future level of pr. year general updates it will be necessary to consider a detailed analysis that would eliminate the distortion causes by errors in the initial registration, or wait new productions to finish.

The experienced gained from the present investigation further suggests that a statistical production monitoring system would benefit future workload estimates.