

Applying DLM and DCM concepts in a multi-scale data environment

Jantien Stoter^{1,2}, Martijn Meijers¹, Peter van Oosterom¹, Dietmar Grunreich³,
Menno-Jan Kraak⁴

¹ OTB, GIS, Technical University of Delft, The Netherlands
{j.e.stoter|p.j.m.vanoosterom|b.m.meijers}@tudelft.nl

² Kadaster, Apeldoorn, The Netherlands
Jantien.stoter@kadaster.nl

³ Bundesamt für Kartographie und Geodäsie (BKG), Germany
dietmar.gruenreich@bkg.bund.de

⁴ ITC, University of Twente, Enschede, The Netherlands
kraak@itc.nl

1. Introduction

This extended abstract presents work in progress in which we explore the DLM and DCM concepts in a multi-scale topographic data environment. The abstract is prepared as input for the Symposium on Generalisation and Data Integration (GDI), University of Colorado, Boulder, 20-22 June 2010.

Brassel and Weibel (1988) and Grunreich (1992) introduced the separation between a Digital Landscape Model (DLM) and Digital Cartographic Model (DCM). The Digital Landscape Model (DLM) contains the basic primary model of reality. DLMs at lower accuracies can be generated by the derivation of primary models at lower semantic and geometric resolution from the basic DLM. In generalisation this is called 'model generalisation'. The Digital Cartographic Model (DCM) is the result of applying cartographic generalisation, i.e. reduction, enlargement, and modification of graphic symbols to the DLMs, see figure 1.

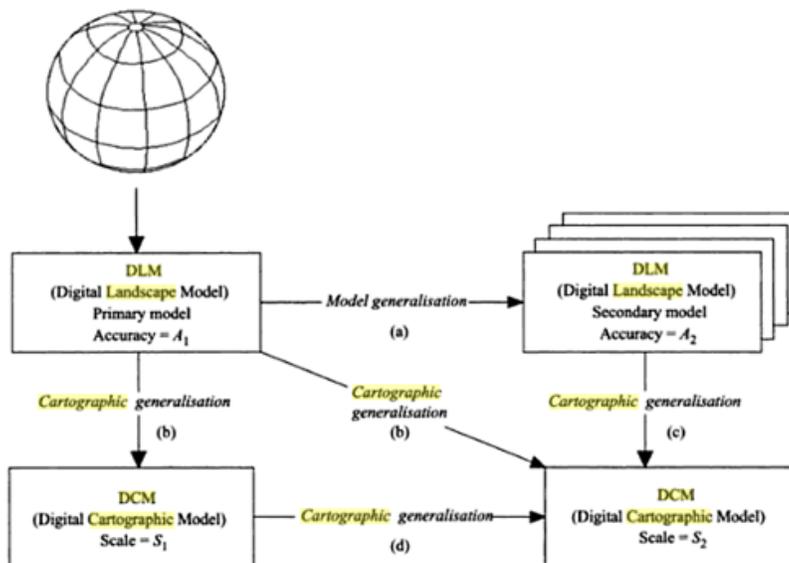


Figure 1: Schematisation of DLM and DCM concept (taken from Joao, 1998)

Although the separation between Digital Landscape Model (DLM) and Digital Cartographic Model (DCM) is considered theoretically as the optimal way of maintaining data sets at multiple scales, in practice data producers, like national mapping agencies (NMA), wrestle with the question what to store explicitly in order to efficiently maintain their geographic databases and maps (Stoter et al., 2010) and consequently what operations to apply to obtain DLMs at lower accuracies and what operations to apply to obtain DCMs.

A main disadvantage of explicit storage of both models, up to the data instance level, is that it leads to more redundancy in multi-scale data models and makes it more difficult to manage geographic databases. This was the motivation of this paper to look in more detail how the DLM-DCM concept works when it is brought in practice. In addition it is interesting to reconsider the DLM-DCM concept now paper maps are no longer the only focus of multi-scale datasets.

Section 2 describes an example of how a NMA applies the DLM-DCM concept by presenting the multi-scale data environment of the Dutch Kadaster (in future work we will also describe the German case). In Section 3 we propose four alternatives of applying the DLM-DCM concept in practice and we evaluate them to conclude on the most sustainable one. Section 4 investigates the vario-scale data structure for the selected solution. Section 5 elaborates on the consequences of DLM-DCM application for data integration as this is the focus of the GDI symposium. The abstract ends with concluding remarks in Section 6.

2. Example of the DLM-DCM concept in practice

Since the eighties the Kadaster produces vector data sets at scales 1:10K, 1:50K, 1:100K, 1:250K, 1:500K and 1:1000K. As the vector data sets originally were generated to support the map making process, they were poorly structured. Since 2007 a TOP10NL dataset, which is the object oriented version of TOP10vector, is available for the whole of the Netherlands. This data set has an improved structure in comparison with TOP10vector. For example it contains planar topology and a river does no longer have a gap because of a crossing road.

A law on key-registers that came into force 1st of January 2010 required also object oriented databases at smaller scales. Since it was expected that it would not be possible to generate these data sets by automated generalisation from TOP10NL, the available vector data sets were converted into object oriented datasets: also here topological errors at crossings were repaired and polygons were split because of crossing lines. Another operation that was applied in the conversion is that crossing infrastructure is identified with line geometry (see figure 2). In TOP10NL roads polygons may be either connecting parts or crossings. In TOP50NL road polygons are collapsed to lines, resulting in crossings with point geometry. In order to have both road crossings and road connections with line geometry, the line segments at crossings were generated as shown in figure 2a. This makes it possible to have different attribute values for road connections and for road crossings.

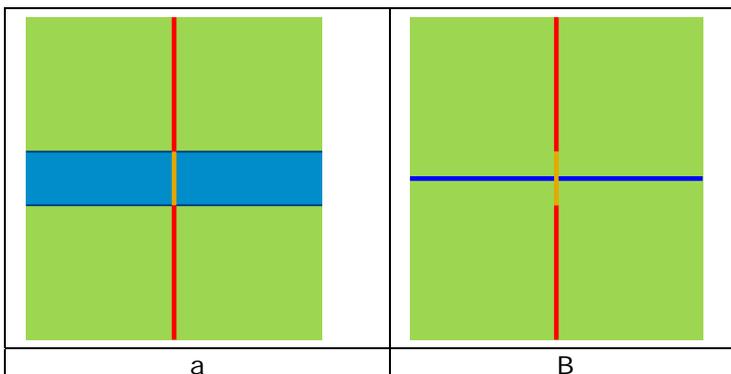


Figure 2: TOP50NL

For the Kadaster case we can say that the DLM-DCM concept is mixed in the smaller scale data sets. The DCM concept is incorporated because the sources of the multi-scale data sets are the TOPxxvector data sets of which the geometries are based on their appearance on the map. For example, a motorway at 1:50k is portrayed with a line-symbol of width 1.5 mm, which is 75 metres in reality. To avoid overlap of the motorway symbol with other features such as buildings, features are displaced and simplified. Creating the map is a simple operation, which adds symbols to the geometries in the vector datasets (see Figure 3).

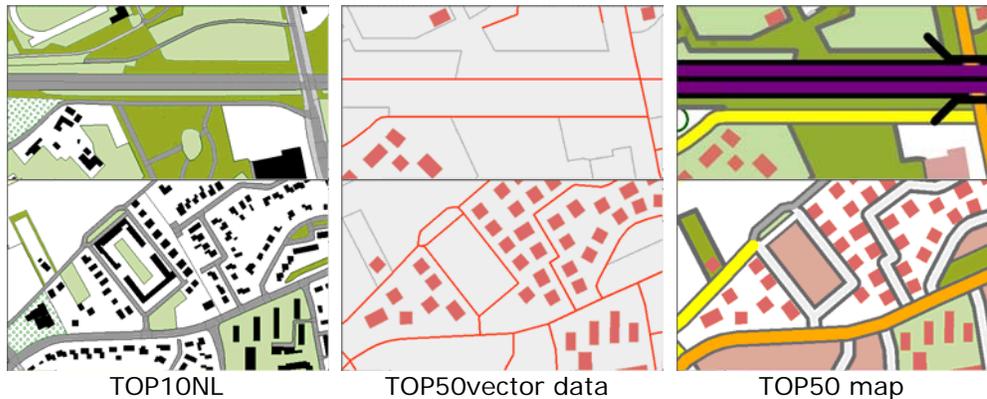


Figure 3: Displacement of features in TOP50vector triggered by graphical conflicts of symbolised features in 1:50k map

In the transition from TOPxxvector to TOPxxNL, we can say that the data set also became a DLM, because of the object creation. As long as the deformations of the data are within the tolerances of the specific scale, the data can still be used in GIS analyses that require less detailed data. However, since the deformations are not controlled, it may be that the inaccuracies due to symbolisation are too big to yield reliable results of some GIS analyses.

3. Where to put the balance between DLM and DCM

Geographic data producers are dealing with data capture, data management and visualisations. For creating a digital geographic database (the DLM), objects will be captured in the real world with certain rules applicable for the data capturing process. Besides rules for geometry and topology, like minimum size, geometric accuracy and connectivity of objects, these rules also include object classification and population of attributes carrying thematic semantics. From the database, objects can be selected to produce digital maps. The transformation of objects from the DLM to a visual end-product is described in a DCM. However, as this transformation process is not in all cases straightforward, data producers face the problem of what to store and maintain, only the geographic objects or also the map objects resulting from this transformation? To make the different choices clear, we evaluate four alternatives of applying the DLM-DCM concept in a multi-scale topographic data environment:

1. It is possible to only store the digital map objects. This is not an optimal solution in the sense that it mixes the representation of 'real world' objects and too much of its visualisation and makes it difficult to use the objects for geographic analysis. An important disadvantage is also that because of the major (and also interactive) adjustments required to solve cartographic conflicts it is hard to maintain datasets at different scales in an integrated manner.

2. One can store the geographic objects persistently (DLMs) and derive the map objects (DCMs) in an automated way when needed. In this case, the geographic objects are thus not adapted **at all** for any kind of visualisation. Although a lot of research has been carried out, a fully automated solution still has not been reached for such a setup. Operations that are difficult to manage in the transformation from geographic objects to map objects are for example displacement and typification.
3. Another option is to store the geographic objects (DLMs) as well as the map objects (DCMs) explicitly. Both models are thus instantiated and made persistent. This allows for fast access to both the geographic objects and the map objects, but comes with a price of redundancy: at every scale both a DLM and a DCM need to be maintained with the accompanying data models. In order to maintain consistency more easily, links can be created between the counterpart instances in both representations. The creation of links and the maintenance of links in order to keep all data sets updated will become a significant task itself.
4. The last option is to store the geographic objects at smaller scales (DLMs) and to adapt them for a default visualisation. The map objects (DCMs) are generated on-the-fly by applying relatively simple visualisation rules. This differs from the second solution that non-straightforward visualisation aspects, e.g. displacements and typification, change the geometry of geographic objects and results of those operations **are** explicitly stored as the geographic objects. This change of geographic objects should however only take place within tolerances, specified in the capturing rules with respect to the desired quality of the dataset. Our motivation to allow this kind of distortion is that other geometric distortions take place within smaller scale datasets anyhow; e.g. simplification, aggregation, or complete removal of objects. An advantage of this approach is that it is easier to establish links between data sets at different scales which better guarantees the consistency. In addition, it provides the possibility to move away from the predefined scale steps that originate from the paper maps and adjust the steps to the digital use of topographic information. A downside of this solution is that problems might still arise when a visualisation is required that significantly differs from the default visualisation. This alternative is partly similar as the solution chosen by the Kadaster. However the distortions in this option are controlled so that they meet the quality criteria of the specific scale, in contrast to the Kadaster solution.

In a multi-scale database the choice how to deal with the geographic objects and map objects is a bit more complicated in comparison when looking at a single scale only. For example, when method three is adapted, both models have to be maintained. In a multi-scale setup this doubles the amount of objects that need to be maintained (per scale the geographic and the map objects). As storage space tends to become cheaper and cheaper this is not a major problem. However, there is also a trend for data producers having to provide higher rates of updates: maintaining more data means more work, so it would be beneficial to minimise the amount of redundancy.

Another problem is that of inconsistencies between the different datasets: if maintenance takes place in a not-fully-automated fashion, it is possible for geographic and map objects to become out of sync with each other (and tell different stories about the same reality).

All in all, the fourth solution we sketched above seems more appropriate: Within the quality bounds required for a geographic dataset, the geographic objects should be adapted to make a default visualisation easily possible. If a user requires higher (geometric or thematic) accuracy, then the multi-scale setup allows to go down the scale dimension to a larger scale dataset and select a more detailed dataset for the application.

Therefore to optimally streamline the process of data production for both analysis and map making purposes, we propose to maintain only the DLM, including additions to apply visualisation rules easier. In the next section we investigate the variable scale data storage to implement this option.

4. Variable-scale as ultimate solution?

From a data producers perspective, the multi-scale setup of the fourth option might still not be ideal. We can extend the line of thinking for the scales stored in the database: If certain features are present at multiple scales, then why store these representations redundantly? Variable-scale data structures have been proposed (Oosterom et al, 2006) and tested (Hofman et al, 2008) to provide an answer. Two advantages of variable-scale data structures are: 1. no, or at least very little, redundancy between scales and 2. also the possibility of 'in-between scales' representations, not only the fixed, stored representations.

Although variable-scale structures might seem to be an optimal solution, it still can be necessary to include a second representation for certain generalisation events for which the representation in the structure (and/or selection at the required scales) will become too complicated. A few examples: 1. It may be required to store both a road area (at the large scales) and a road centerline (at the smaller) scales and link these representations in the structure to the same real world object, 2. Certain generalisation operations require contextual information and are relatively expensive to compute; in these cases it may be more effective to store a second representation, such as a displaced house, and 3. certain types of concepts occur not at the largest scale, but only at smaller scales (e.g. roundabout composed by several road areas, or building block composed by individual houses and optionally gardens).

In current map products the decisions when to add extra representations are 'black-white' and bound to the maintained scales: road areas are present on 1:1,000 and road areas and roundabouts on 1:10,000, individual houses on 1:1,000, individual houses and building blocks on 1:10,000 and only building blocks in urban areas on 1:50,000.

In further research we will investigate how to embed these semantic aspects in the variable-scale data structures. The nature of objects changes in a continuous way related to geographic scale, and thus there is no reason to take only black-white decisions. It may be far more natural to gradually move from individual houses to building blocks when moving from large to smaller map scales (and thus being able to provide smooth zooming to end users).

5. Consequences for data integration

Data integration within the context of DLM-DCM concepts in a multi-scale data environment has three meanings: integration of data within one scale, integration of data between different scales and integration of the multi-scale data with external data sources.

1. Data integration within one scale

Topography is an integration of several themes (water, buildings, roads) that have been identified as separate themes within the European INSPIRE Directive (INSPIRE, 2007). The integration can be accomplished in two ways. The first option is to integrate the independently stored themes. The second option is to maintain the themes in an integrated DLM-DCM structures representing separate scales while fully acknowledging the concepts as defined by INSPIRE. The separate themes as required by INSPIRE can then be generated from this structure. It is obvious that the second option enforces consistency between themes and is therefore better for maintaining the data (but because of different data owners not easy to realise).

2. Data integration between scales

Data at separate scales need to be integrated. This can be accomplished within a vario-scale representation as described above and as shown in Oosterom et al (2006).

3. Data integration with external sources

Finally multi-scale topographic data should be prepared to integrate with external data sources, i.e. to link with similar objects in external data sets at the appropriate scale. The main advantage of the integrated approach of the vario-scale is that external data has to be linked only once instead of linking to topographic data sets at multiple scales. How to deal with the linked objects within the other scales (for example display, analysis) is a nice topic for future research.

6. Concluding remarks

This abstract, which presents work in progress, elaborated on how the DLM-DCM principles work when they are applied in practice by first presenting a case study from the Netherlands. In future work we will include more practices such as Germany, Denmark, Sweden.

In this abstract we also presented four alternatives to balance DLM and DCM in practice. We concluded that the optimal alternative for maintaining and providing multi-scale information is the alternative where geographic objects at smaller scales (DLMS) are adapted for a default visualisation. The corresponding DCMs may be generated on-the-fly by applying relatively simple visualisation rules. Future research will further study this approach. Topics to be addressed are:

- What are the specifications for the integrated DLM-DCM multi-scale representations?
- How does the theme integration within one scale and across multiple scales works in the context of INSPIRE?
- How can the vario-scale structure support the integrated DLM-DCM multi-scale representations?
- How to integrate 3D and time in this structure (see Oosterom and Stoter, 2010)?

References

Gruenreich, D. (1992) ATKIS - a topographic information system as a basis for GIS and digital cartography in Germany. From digital map series to geo-information systems, Geologisches Jahrbuch Reihe A (Heft 122), pp.207 - 216

Hofman, A., Dilo, A., Oosterom, P. J. M. van, and Borkens, N. (2008). Using the constrained tGAP for generalisation of IMGeo to Top10NL model. In Mustiere, S., Mackaness, W., and Oosterom, P. J. M. van, editors, Proceedings of the ICA Commission on Map Generalisation and Multiple Representation, pages 1-23, Montpellier.

INSPIRE (2007), <http://inspire.jrc.ec.europa.eu/>.

João, E. (1998) Causes and Consequences of Map Generalisation. London: Taylor & Francis.

Oosterom, P., van, Vries, M. E. de, and Meijers, B. M. (2006). Vario-scale data server in a web service context. In Ruas, A. and Mackaness, W., editors, Proceedings of the ICA Commission on Map Generalisation and Multiple Representation, pages 1-14, Paris.

Oosterom, P.J.M., van and J.E. Stoter (2010). 5D Data Modelling: Full Integration of 2D/3D Space, Time and Scale Dimensions. Accepted for GIScience, Zurich, September 2010.

Stoter, J.E., P.J.M. van Oosterom, C.W. Quak, T. Visser, N. Bakker (2010), A semantic rich Multi-Scale Information Model Topography, accepted for publication in International journal of geographical information science (IJGIS).

Weibel, R., and G. Dutton (1998) 'Constraint-Based Automated Map Generalization' Proceedings of the 8th Spatial Data Handling Symposium, Vancouver, pp. 214-224.