

SEMANTIC 3D CITY MODEL TO RASTER GENERALISATION FOR WATER RUN-OFF MODELLING

Edward Verbree ^a, Marian de Vries ^a, Ben Gorte ^b, Sander Oude Elberink ^c, Gosro Karimlou ^d

^a Delft University of Technology, OTB Research, Faculty of Architecture and the Built Environment, the Netherlands
(e.verbree, m.e.devries)@tudelft.nl

^b Delft University of Technology, Faculty of Civil Engineering and Geosciences, the Netherlands
b.g.h.gorte@tudelft.nl

^c University of Twente, Faculty of Geo-Information Science and Earth Observation, the Netherlands
s.j.oudeelberink@utwente.nl

^d Hydrologic, Amersfoort, the Netherlands
gosro.karimlou@hydrologic.com

3D GeoInfo 2013 – 8th 3D GeoInfo Conference
ISPRS WG II/2 Workshop

KEY WORDS: Multi-scale modelling, water run-off modelling, 3D city models, semantic generalisation

ABSTRACT:

Water run-off modelling applied within urban areas requires an appropriate detailed surface model represented by a raster height grid. Accurate simulations at this scale level have to take into account small but important water barriers and flow channels given by the large-scale map definitions of buildings, street infrastructure, and other terrain objects. Thus, these 3D features have to be rasterised such that each cell represents the height of the object class as good as possible given the cell size limitations. Small grid cells will result in realistic run-off modelling but with unacceptable computation times; larger grid cells with averaged height values will result in less realistic run-off modelling but fast computation times. This paper introduces a height grid generalisation approach in which the surface characteristics that most influence the water run-off flow are preserved. The first step is to create a detailed surface model (1:1.000), combining high-density laser data with a detailed topographic base map. The topographic map objects are triangulated to a set of TIN-objects by taking into account the semantics of the different map object classes. These TIN objects are then rasterised to two grids with a 0.5m cell-spacing: one grid for the object class labels and the other for the TIN-interpolated height values. The next step is to generalise both raster grids to a lower resolution using a procedure that considers the class label of each cell and that of its neighbours. The results of this approach are tested and validated by water run-off model runs for different cell-spaced height grids at a pilot area in Amersfoort (the Netherlands). Two national datasets were used in this study: the large scale Topographic Base map (BGT, map scale 1:1.000), and the National height model of the Netherlands AHN2 (10 points per square meter on average). Comparison between the original AHN2 height grid and the semantically enriched and then generalised height grids shows that water barriers are better preserved with the new method. This research confirms the idea that topographical information, mainly the boundary locations and object classes, can enrich the height grid for this hydrological application.

1. INTRODUCTION

1.1 HydroCity: key to flood-resilient cities

Floods caused by excessive rainfall have disastrous effects on many cities around the world. Flooding-induced disasters are becoming increasingly frequent and range from local, water-related inconvenience to flood-induced disruption of society and devastation. HydroCity (hydrocity, 2013) is a platform where several knowledge institutes and Dutch private and public organisations in the water and earth-observation sectors have joined forces to help cities improve their flood resilience.

HydroCity covers the full workflow to improve flood resilience (Figure 1): data acquisition and storage, data analysis and modelling, and decision-support applications (Dassen, 2012a, 2012b).

This paper describes the creation of a detailed digital surface model represented by a raster grid at several appropriate cell spacing following a semantic generalisation procedure that preserves the surface characteristics relevant to water run-off prediction. This approach is tested and validated by the CityFlood water run-off model to confirm the idea that topographical information, mainly the boundary locations and object classes, can enrich the height grid for this hydrological application.

The outline of this paper is as follows. First we describe the basic idea of the combination of high-density laser data with a large-scale object-based topographic map into a TIN-based digital surface model (DSM). In this study we have used the National height model of the Netherlands AHN2 and the large scale Topographic Base map (BGT, map scale 1:1.000). Then we discuss the generation of two grids with 0.5m cell spacing: one grid for the object class labels and the other for the TIN-

interpolated height values. Hereafter the procedure to generalise these grids to sparser cell sizes is described, in which the semantics of the object classes are taken into account in the generalized height grid mapping. Finally the run-off modelling test results are given.



Figure 1 Water surface flow in the city

2. DIGITAL SURFACE MODEL

2.1 Combination

The basic idea is that we combine high density laser data (e.g. AHN2) with a large scale object based topographic map (e.g. BGT). See Figure 2.

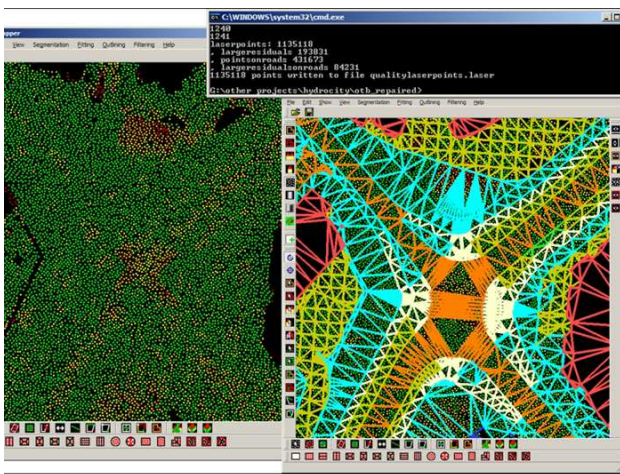


Figure 2 Overlay of BGT and laser points

The assumption is that these two datasets contain enough information to generate a detailed 3D topographic map for hydrological purposes. The main requirements are:

- Water flows from high to low; it is important to capture height differences;
- Small height differences at street level should be kept in the final 3D model;
- The shape of main buildings as such is not so important as the water will run into the sewage system; In CityFlood water run-off modelling the buildings are left out of the grid, by putting a NaN at those grid cells;
- 3D map should be converted to a grid structure in order to act as input for hydrodynamic software, such as CityFlood.

Above mentioned requirements are translated in the following implications:

- The 2d map should contain information on the class or function of every object;
- The neighbours of every object should be known (including their class or function);
- The map should be geometrically and semantically accurate at the highest level-of-detail that is available.

After fusion of the two data sets the program knows which points belong to which polygon, including the label of that polygon and vice versa, see Figure 3.

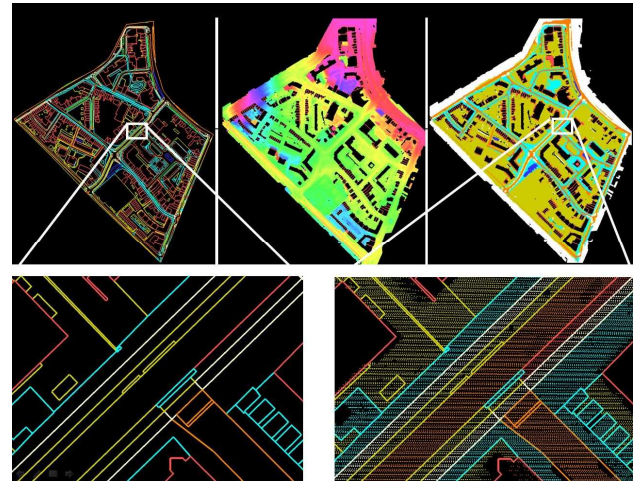


Figure 3 Polygons fused with airborne laser data

The conventional way of road modelling is to create a TIN structure from boundaries from one road side to the other. In the HydroCity project it was found that it is wise to check whether the road can be generalised like that, or whether additional points should be inserted at the road surface, similar to what is done at terrain surfaces. We implemented two options: in a regular grid (say every 3x3 meter) or only if that point makes sense (mesh simplification). If we add these grid points, the number of large residuals decreases enormously.

2.2 Set of rules

In the following the assumptions and rules are specified per topographic class. It describes the steps to select and process the laser data. It is important to remember that the 2D map is delivering the 2D locations of all boundaries, the semantics of constraints between neighbouring polygons and constraints within each polygon. The Lidar data is processed according to the rules of its corresponding polygon. In that sense the map data is more 'active' or 'leading' than the laser data.

Polygon based rules:

Class 'Water': the surface is determined by triangulation of boundary object points; all ground points are set to average height, to represent a horizontal plane.

Class 'Roads': lidar points are inserted inside polygon, interpolated to a regular grid, mesh simplification to reduce number of points, followed by constrained triangulation; each object point is determined by height of local fitted plane.

Class: 'Terrain': lidar points are inserted inside polygon, interpolated to a regular grid, mesh simplification to reduce

number of points, followed by constrained triangulation; each object point is determined by height of local fitted plane.

Class ‘Buildings’: the surface is determined by triangulation of boundary points; all object points are set to average height.

Neighbouring polygons at boundary locations do have their own set of rules. For example: where ‘buildings’ touches another class the objects will keep their own height, a vertical wall will be created in-between. Where ‘terrain’ touches ‘Water’, ‘Road’, or ‘Terrain’ the boundary will take the height of the other class.

2.3 Case Amersfoort

After applying the rules to the dataset of Amersfoort, the following models can be made, see Figure 4.

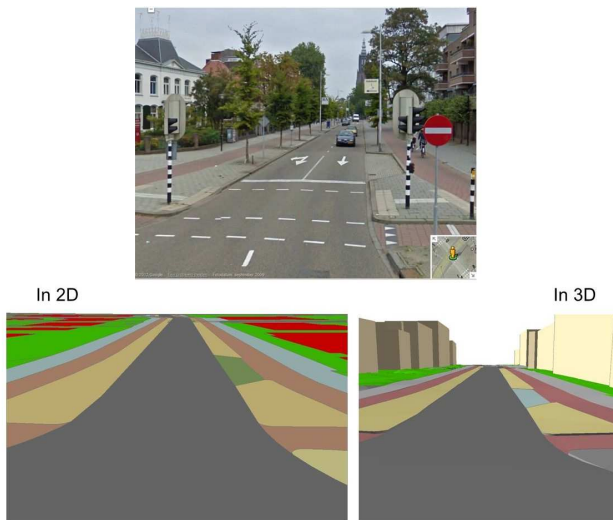


Figure 4 2D and 3D View of the Utrechtsestraat in Amersfoort

2.4 From TIN to Raster

An artificial grid is generated with gridsize 0.5 meter. At each grid location, a point in polygon operation selects the corresponding triangle from the 3D TIN model. The height at the grid location is taken as raster cell height, the label of that triangle is stored in a label raster. If the triangle is from a building, the raster cell height value is set to -99 (NaN), as these are not being part of the run off model in CityFlood.

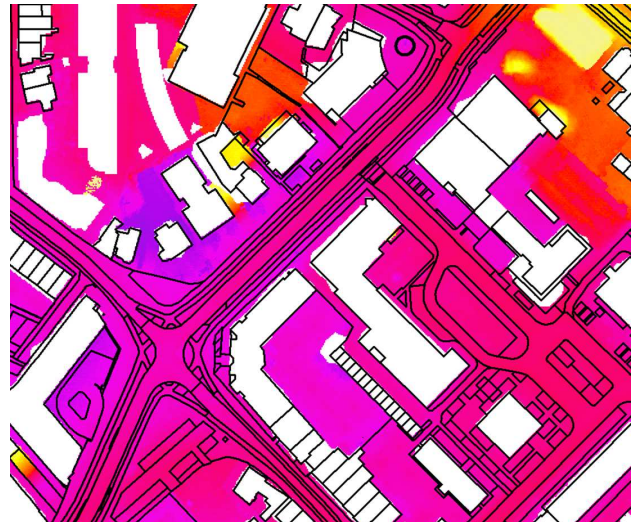


Figure 5 Height Grid – interpolated with semantics

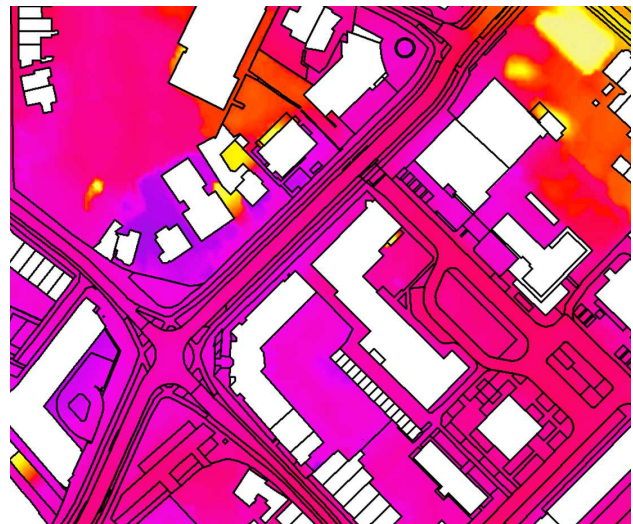


Figure 6 Height Grid – interpolated without semantics

The differences with a ‘non-topographical’ grid (Figure 5 versus Figure 6) are:

- The first difference is the way how laser points are selected as input for the raster cell height calculation. The normal DTM takes just nearby laser points, the hydro TIN is based on points within that polygon. This difference only will occur at places where there is more than one polygon within a grid cell (so at boundary locations);
- Second difference is the height determination itself: a weighted (inverse distance) interpolation versus a plane fitting through the selected points. Consequences are that the noise is a bit higher in normal DTM (as it mainly depends on noise of nearest laser point) and that there can be a small difference if the area within the grid cell is not planar;
- Third difference is the semantics that are incorporated in the hydro TIN. The semantics deliver the (hard) constraints to the possible shape of objects, e.g. a water object should be horizontal, or a terrain object should glue to a neighbouring road or water object. In case (some of) the laser points do not fulfil these constraints, a difference between a normal DTM and a hydro TIN is caused.

The following step is to generalise this grid by making the pixels twice as large.

3. RASTER GENERALISATION

The objective in this section is to generalize a DSM, i.e. to interpolate heights at a lower resolution, to be used for run-off modelling.

As a result of runoff modelling the area gets subdivided into catchments, sub-catchments, etc.. In fact, every pixel belongs to a sub-catchment, being the set of pixels that eventually drain into that pixel. The purpose of generalization is to represent the area in fewer pixels in such a way that the result of run-off modelling, in terms of generated catchments, changes as little as possible. When analyzing run-off results of the original and the generalized DSM, the same pattern should become apparent. For example, during the run-off modelling process “streams” are usually being formed, through which eventually large amounts of water will be flowing, where as elsewhere barriers are preventing water to go from one area (i.e. sub-catchment) to another. We require that during generalization the important streams do not get blocked, and the important barriers do not get broken. Unfortunately it is difficult to recognize streams and barriers, and especially their importance, without running the model first. This, however, would violate the purpose of the entire generalization, which was to speed up run-off modelling. Therefore we propose to use semantics as an alternative information source for recognizing which topographic objects potentially act as streams, and which ones as barriers. The semantics are derived from the class label grid, which is created alongside the elevation grid (DSM). Subsequently we may try to preserve those objects and their role in the model. As preserving certain objects might go at the expense of other objects priorities have to be introduced into the process.

3.1 Influence of priorities

When objects are large compared to the pixel size, and represent areas rather than being linear (also this has to be seen in relation to the pixel size), then it is not necessary that some objects would have to get priority over others.

Priorities get important when objects are getting longer and narrower. At the left of Figure 7 you see a stream (in black). Provided that the elevations are arranged correctly, it is a valid stream in a run-off model: from each pixel the water will flow to the lowest of the 8 neighbours. In the centre you see in white what might be a dike. Even when all its pixels are high enough, however, it is not a valid dike: water can diagonally flow through at many places. Dikes (and barriers in general) need to be 4-connected. The dike at the right is a valid one.

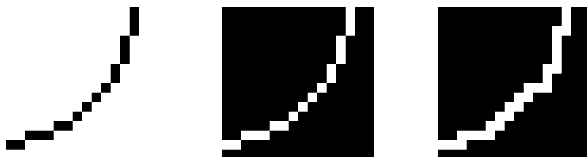


Figure 7 Stream (valid), Dike (invalid), Dike (valid)

The property that barriers must be 4-connected, whereas streams are 8-connected should still be satisfied after generalization. Moreover, for area objects generalization should maintain the (relative) sizes. Therefore, semantics are used to identify stream and barrier classes, allowing different generalization strategies for both. Those strategies should be

applied only when objects, belonging to those classes, are linear.

Generalization gets hard in case of conflicts. Probably there are many cases, but an obvious one in the context of the above is shown in Figure 8 at the left.

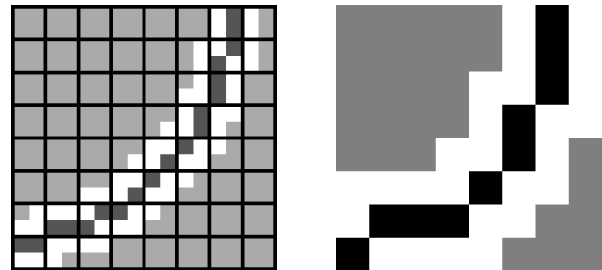


Figure 8 Original stream/dykes; required generalization

It is a typical Dutch stream surrounded by two dikes. Like before, the stream is 8-connected, and the dikes are 4-connected. The required result is shown in Figure 8 at the right. To generalize the stream and the dike as indicate require to know a lot of semantics about the area: some pixels in the generalization should become 'dike', where there is no dike at all in the corresponding 2x2.

3.2 Case Amersfoort

An experiment was executed in order to assess the generalization method described above on real data. As input for the experiment the two raster data sets were used, as derived at the highest resolution of 0.50m from the HydroTIN of Amersfoort: class label and height (see section 2.3). The task was step-wise generalisation to 1, 2 and 4 m resolution.

During each generalization step different strategies were used for object classes 'area' (buildings and terrain), 'stream' (roads) and 'barrier' (speed humps), as described above, on the basis of the land-cover raster map. Four binary (0/1) maps were created and generalized separately according to their respective strategies to the lower resolutions.

At each target resolution the four binary maps are combined into a single “type” map with values 1-4 (barriers, streams, buildings and terrain), and 0 outside the AOI. Ideally this would exactly fill the entire map, but as already announced, conflicts may arise (several types at one pixel), for which a descending priority order is applied (1=highest to 4=lowest). On the other hand, a few holes remain (pixels getting no value at all); these were set to 4 (terrain).

To reconstruct, at any generalized resolution, the land cover map from the type map, at each pixel the corresponding window at the highest resolution is examined: of pixels belonging to the same type, the majority is chosen, or the lowest number in case of a draw.

The result of this approach from the original resolution, see Figure 9, to its 4mx4m generalization land cover map is shown in Figure 10.



Figure 9 Land Cover – original resolution 0.50m

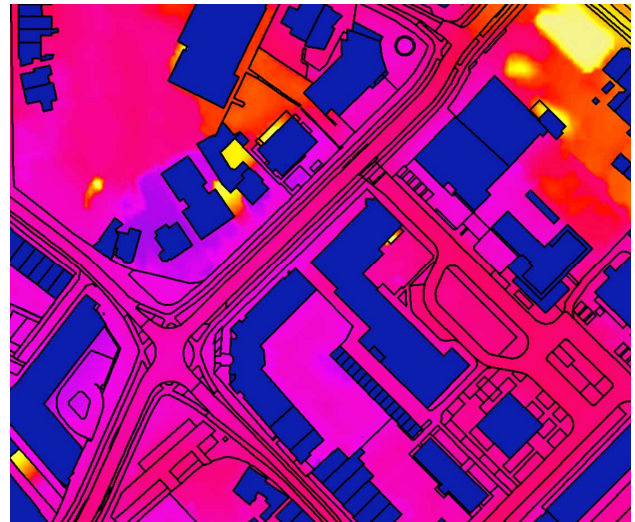


Figure 11 Height - original resolution 0.50m

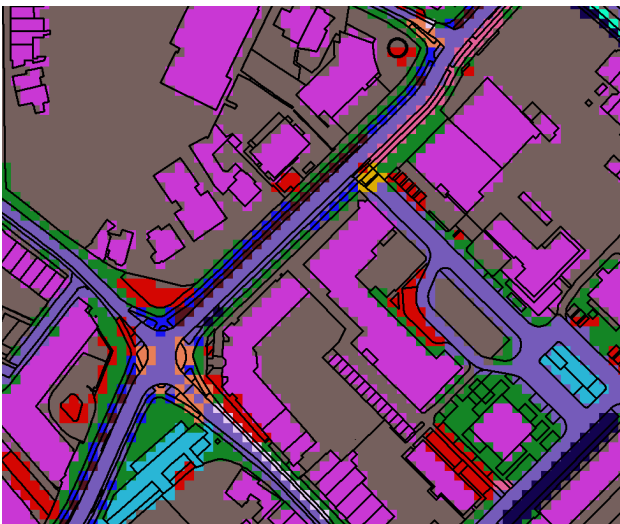


Figure 10 Land Cover - generalized resolution 4m

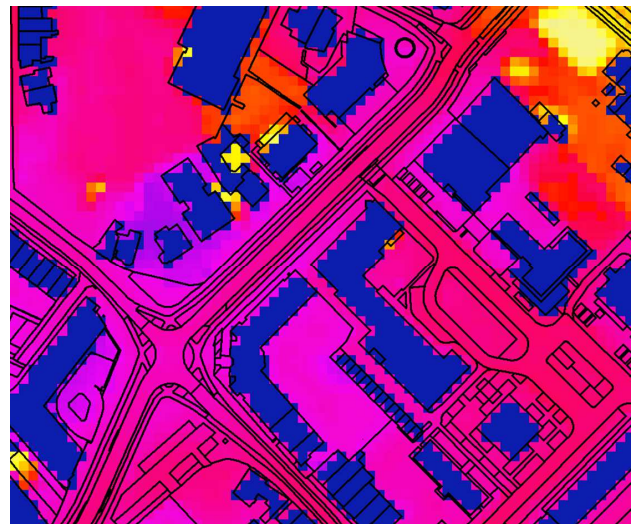


Figure 12 Height - Generalized resolution 4m

The height values of the pixels at the highest resolution, belonging to the same type as a generalized pixel at a lower resolution, are averaged into a generalized height. Here the distinction between building and terrain is important: the height of a generalized terrain pixel is determined only by terrain heights and not contaminated by the building.

The result of this approach from the original resolution, see Figure 11, to its 4mx4m generalization height map is shown in Figure 12.

4. CITYFLOOD MODELLING

The results of this approach are tested and validated by water run-off model runs with the CityFlood software for different cell-spaced height grids. Input of this model is on the one hand the land cover and height grids, and historical, fictive or forecasted rainfall data on the other hand. The output shows dynamically to which sink (open soil, sewer, surface water) this water flows. If the rainfall intensity exceeds the infiltration capacity, run-off will occur. A visual comparison gives an indication whether the water follows the intended paths to the sinks, or whether it floods onto the non-intended surfaces (roads).

This detailed view is now possible because the model uses the semantic processed and generalized height and land cover grids. The output of one of the model runs is shown in Figure 13 (red: water at street; orange: water in gullies; yellow: water in sewer system; green: interception; blue: infiltration).

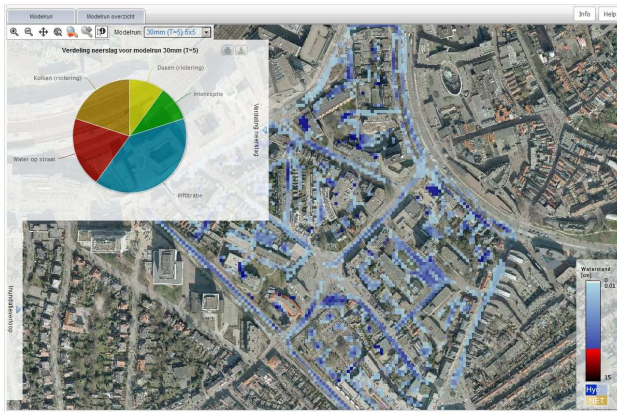


Figure 13 CityFlood output

5. CONCLUSIONS

Water run-off modelling applied within urban areas requires an appropriate detailed surface model represented by a raster height grid. Accurate simulations at this scale level have to take into account small but important water barriers and flow channels given by the large scale map definitions of buildings, street infrastructure, and other terrain objects. This semantic based gridding causes some differences with a conventional, non-topographical grid. The first difference is the way how laser points are selected as input for the raster cell height calculation. The second difference is the height determination itself: a weighted (inverse distance) interpolation versus a plane fitting through the selected points. Third difference is the semantics that are incorporated in the hydro TIN. The semantics deliver the (hard) constraints to the possible shape of objects, e.g. a water object should be horizontal, or a terrain object should glue to a neighbouring road or water object.

The purpose of generalization is to represent the area in fewer pixels in such a way that the result of run-off modelling, in terms of generated catchments, changes as little as possible. Semantics are used as an alternative information source for recognizing which topographic objects potentially act as streams, and which ones as barriers. The semantics are derived from the class label grid, which was created alongside the elevation grid. Subsequently we try to preserve those objects and their role in the model.

Thus, topographical information, mainly the boundary locations and object classes, can enrich the height grid for hydrological applications.

REFERENCES

Dassen, W., Loovers, I., Lobbrecht, A., Doelmatig stedelijk waterbeheer in Amersfoort, 2012, H₂O Magazine, 14/15.

Dassen, W.H., Havers, B., Lobbrecht, A.H., Loovers, I., 2012, HydroCity - integrated water information services for city management, International Water Conferences, 4-6 November 2012, Amsterdam, The Netherlands

HydroCity, www.hydrocity.com, accessed June 2013