Improved methodology for processing raw LiDAR data to support urban flood modelling – accounting for elevated roads and bridges

A. F. Abdullah, Z. Vojinovic, R. K. Price and N. A. A. Aziz

ABSTRACT

Digital Terrain Models (DTMs) represent an essential source of information that can allow the behaviour of the urban floodplain, and its interactions with the drainage system, to be examined, understood and predicted. Typically, such data are obtained via Light Detection and Ranging (LiDAR). If a DTM does not contain adequate representation of urban features the results from the modelling efforts can be. This is due to the fact that urban environments contain variety of features, which can have functions of storing and/or diverting flows during flood events. The work described in this paper concerns further improvements of a LiDAR filtering algorithm which was discussed in a previous work. The key characteristics of this improved algorithm are: ability to deal with buildings, detect elevated road and represent them accordance to reality and deal with bridges and riverbanks. The algorithm was tested using a real-life data from a case study of Kuala Lumpur. The results have shown that the newly developed MPMA2 algorithm has better capabilities of identifying some of the features that are vital for urban flood modelling applications than any of the currently available algorithms and it leads to better agreement between simulated and observed flood depths and flood extents.

Key words | Digital Terrain Models (DTMs), LiDAR filtering algorithms, urban features, urban flood modelling

A. F. Abdullah (corresponding author)
Z. Vojinovic
R. K. Price
Department of Hydroinformatics and Knowledge Management,
UNESCO-IHE,
Westvest 7,
NL-2611 AX Delft,
The Netherlands
E-mail: *abdul42@unesco-ihe.org*

N. A. A. Aziz
Dr. Nik & Associates Sdn. Bhd.,
No 22 and 24, Jalan Wangsa Delima 6,
KLSC, Seksyen 5,
Pusat Bandar Wangsa Maju,
53300 Kuala Lumpur,
Malaysia

INTRODUCTION

The role of modelling within urban flood management is in complementing the acquisition of data to improve the information and understanding about the performance of a given drainage network, taking into account the associated urban terrain. Considerable attention has been given to the acquisition of good geometric and topographical data at adequate resolution in order to describe the primary features of the flow paths through the urban area. In this respect, a Digital Terrain Model (DTM) represents one of the most essential sources of information that is needed by flood managers. A DTM refers to a topographical map, which contains terrain elevations and, as such, it is used to characterize the terrain (or land) surface and its properties. It is a representation of the Earth's surface (or a subset of it) and it strictly excludes features such as vegetation, buildings, bridges, etc.

In urban flood modelling, DTMs are required for the analysis of the terrain topography, identification of overland flow paths and for setting up 2D hydraulic models. Nowadays, one of the most preferred techniques for modelling floods in urban areas is a coupled 1D/2D modelling approach. This technique can be used to describe the dynamics and interaction between surface and sub-surface systems. For an efficient use of 2D models the collection and processing of terrain data is of vital importance. Typically, Light Detection and Ranging (LiDAR) surveys enable the capture of spot heights at a spacing of 0.5 m to 5 m with a horizontal accuracy of 0.3 m and a vertical accuracy of 0.15 m. Most LiDAR surveys result in a substantial amount of data, which requires careful processing before it can be used for any application. Recently, the vertical

doi: 10.2166/hydro.2011.009

accuracy of LiDAR has increased dramatically to 0.05 m (FLI-MAP 2010). Thinning, filtering and interpolation are techniques that are used in the processing of LiDAR data.

Filtering is a process of automatic detection and interpretation of bare earth and objects from the point cloud of LiDAR data, which results in the generation of a DTM. To date, many filtering algorithms have been developed, and in a more general sense, most of them have become standard industry practice. However, when it comes to the use of a DTM for urban flood modelling applications these algorithms cannot always be considered suitable and, depending on the terrain characteristics, they can even cause misleading results and degrade the predictive capability of the modelling technique. This is largely due to the fact that urban environments often contain a variety of features (or objects), which have the functions of storing or diverting flows during flood events. As these objects dominate urban surfaces appropriate filtering methods need to be applied in order to identify such objects and to represent them correctly within a DTM so it can be used more safely in modelling applications. However, most of the current filtering algorithms are designed to detect vegetation and freestanding buildings only and features such as roads, curbs, elevated roads, bridges, rivers and river banks are always difficult to detect. Therefore, further improvements of LiDAR filtering techniques are needed so that modelling efforts can generate more fruitful results. The work described in this paper provides a contribution in this direction as it deals with some of these important issues (i.e. representation of elevated roads and bridges). It represents a further expansion of the work described in Abdullah *et al.* (2011).

URBAN FLOOD MODELLING

Typically, urban flood modelling practice concerns the use of 1D, 1D/1D, 2D and 1D/2D modelling approaches; see, for example, Chen *et al.* (2005), Garcia-Navarro & Brufau (2006), Hunter *et al.* (2007, 2008), Kuiry *et al.* (2010) and Price & Vojinovic (2011). Mark *et al.* (2004) demonstrated how the 1D modelling approach can be used to incorporate interactions between (i) the buried pipe system, (ii) the streets (with open channel flow) and (iii) the areas flooded with stagnant water. Djordjevic *et al.* (2005) have implemented a dual drainage concept (which represents a combination of minor and major systems) within a 1D model. Vojinovic & Tutulic (2009) have explored the difference in predictive capabilities of 1D and 1D/2D modelling approaches for the purpose of urban flood analysis across irregular terrains and their corresponding damage estimation. Also, Vojinovic *et al.* (2011) have shown how different terrain data resolutions, features such as roads and building structures, and different friction coefficients can affect the simulation results of 1D/2D models.

The literature to date confirms that, apart from different model formulations, the variations in the ground topography, discontinuities, representation of features, surface roughness and terrain data resolution are important factors that need to be carefully considered and accounted for in the flood modelling studies.

Typically, 1D models are used to simulate flow through pipes, channels, culverts and other defined geometries. The system of 1D cross-sectional-averaged Saint-Venant equations which are used to describe the evolution of the water depth and either the discharge or the mean flow velocity consist of the conservation of mass (continuity equation) and momentum. The boundary conditions are either discharges or water levels (or depths) at conduit/channel ends. In a channel network, these are not known in advance and need to be determined by the numerical solution procedure. The solution is commonly based on a temporary elimination of variables at internal cross sections and the reduction of all equations to a system of unknown water levels.

The system of 2D shallow water equations consists of three equations: one continuity and two equations for the conservation of momentum in two Cartesian coordinates. The simulation process in the case of coupled 1D–2D modelling is based on complex numerical solution schemes for the computation of water levels, discharges and velocities. The surface model (i.e. 2D model) simulates vertically integrated two-dimensional unsteady flow given the relevant boundary conditions (e.g. resistance coefficients, etc.) and bathymetry (as provided by a DTM of the catchment area). The interactions between channels and floodplains are determined according to the type of link. For example, discharges generated by pumping stations, weirs or orifices are regarded as the lateral inflow to the 2D model. Also, if the channel flow exceeds the ground level (for pipe network systems) or

bank levels (for open channels) then the discharge is computed by the weir (or orifice) discharge equation and it is considered as the lateral inflow to the 2D model.

The two domains (1D and 2D) are normally coupled at grid cells overlying the channel computational points through mutual points of the connected cell and the adjoining channel section (Vojinovic & Tutulic 2009; Price & Vojinovic 2011). The dynamic link that allows for the interaction of mass and momentum fluxes between the two model domains has been implemented in several commercial software packages. Such examples are MIKE 11 for the 1D modelling system and MIKE 21 for the 2D modelling. Solving for water flows on a regular grid, as in the case of MIKE 21, has the advantage of providing an easy integration with DTMs which are most often available in a regular grid format. In the work described in this paper, the modelling part was carried out by the 1D/2D modelling technique using the above-mentioned commercial packages: MIKE 11 and MIKE 21.

ISSUES CONCERNING SPATIAL DATA AND FLOOD PROPAGATION

Urban features

Apart from the variations in the ground topography and surface roughness the flow of flood water in urban areas is also controlled by the presence and distribution of buildings, roads, elevated roads and other objects. In this respect, remote sensing technology and the resulting geospatial data richness now afford a detailed description of urban landscapes and opportunities exist to improve urban flood models by tailoring schemes to this data. In this respect, the LiDAR technology has been invaluable. The main product from a LiDAR survey is a Digital Surface Model (DSM) consisting of 3D coordinates defining the ground surface and above-surface features, such as vegetation, buildings and vehicles, that return a signal from the laser pulse. To realistically account for floodplain storage and conveyance, multidimensional flood models require a terrain height representative of (relatively) impervious boundaries such as bare-earth and solid walls. Hence, filtering of the DSM is necessary to remove unwanted above-ground features to leave a bare-earth model or DTM. Once the unwanted features are removed, the above-ground features that are important for flood modelling can then be added back onto the DTM surface. As mentioned above, in urban scenes non-terrain objects, such as buildings and walls, can have a strong influence on water flow. When these features are poorly captured by LiDAR data they can then be inserted into the DTM either manually or automatically in the case that automatic routines exist (Schubert *et al.* 2008).

Elevated roads and bridges

Elevated roads and bridges are common objects in urban areas and their purpose is to span the gap between two land masses. Such objects, and particularly elevated roads, can cause unrealistic obstruction to the flow of water. In some cities around the world, and particularly in Asia (e.g. Kuala Lumpur, Bangkok, Tokyo, etc.) elevated roads are very common and they can occupy a significant portion of urban space. Current literature concerning LiDAR data processing shows very little progress in detecting such structures and most of the work to date is based on the use of satellite imagery and radar. Evans (2008) utilized a 'fuzzy'-based approach for the identification of the bridge edges with a limited amount of available LiDAR data. This methodology allowed for the detection of areas that are deemed likely to be bridge edges. There are also several studies where Synthetic Aperture Radar (SAR) has been used to detect bridges in urban areas (see, for example, Houzelle & Giraudon (1992) and Wang & Zheng (1998)). The principle used in the detection of bridges is based on the assumption that every bridge contains the same material as nearby roads. Trias-Sanz & Lomenie (2003) proposed an artificial neural-network-based approach which can detect bridges in high resolution satellite imagery data using the following assumptions:

1. Large regions of water or railway yards are separated by a narrow and long strip. This strip is a bridge.
2. A small gap between two regions that have been identified as roads and are aligned is a bridge. The definition of 'small' is defined by the user.
3. A small gap between two regions that have been identified as canal and are aligned is a bridge. The definition of 'small' is user-defined.

Other authors have used a segmentation method to process LiDAR data. Sithole & Vosselman (2006) have used a segmentation method to detect bridges under the following assumptions:

1. A bridge is connected to the bare earth on at least two sides.
2. To span land masses a bridge will, along its length, necessarily is raised above the bare earth. Hence, along the length of a bridge, diametrically opposite points on its perimeter are raised above the bare earth.
3. A bridge is typically built to be longer than its width.

In order to properly represent elevated roads and bridges in a DTM that is required for 1D/2D flood modelling purposes, two different approaches can be applied. For a normal bridge (i.e. a bridge that crosses a river), it should be completely removed from the DTM as it is normally accounted for in a 1D model. For an elevated road, while this object needs to be removed from the DTM, if the structures underneath (such as piers) occupy larger areas which may cause diversion of the flow, they may still be needed. Currently available filtering algorithms can only partially remove an elevated road from the DTM. In such a case, the algorithm would leave behind a non-uniform topography that generates an obstruction that does not necessarily correctly represent the reality under the elevated roads. Furthermore, besides the major structure, there may be structures underneath (such as piles) that need to be considered. As these (lower) structures can play an important role in description of a flood pattern it means that careful attention should be paid to their treatment and representation.

Processing of LiDAR data

LiDAR is an active sensor (similar to radar), which transmits laser pulses at a target and records the time it takes for the pulse to return to the sensor's receiver. This technology is currently being used for high-resolution topographical mapping, by mounting a LiDAR sensor, integrated with a Global Positioning System (GPS) and Inertial Measurement Unit (IMU) technology, to the bottom of an aircraft and measuring the pulse return rate to determine surface elevations. In many cases, LiDAR is relatively less expensive, faster and more accurate than conventional photogrammetric methods of topographical mapping. In addition, the data derived from LiDAR is digital and therefore ready to be used in a GIS package. In recent years, many researchers and practitioners have become more familiar with the advantages of this technology and there are many efforts that focus on maximizing the use of LiDAR data.

This technique delivers not only detailed information about the geometry, but also about the reflectance characteristics of the Earth's surface in the laser wavelength, which is typically in the near-infrared (NIR) spectra between 800 nm and 1,550 nm wavelengths. It has been reported that most LiDAR systems currently record the return amplitude of each received echo and a growing number of systems already provide full-waveform digitization (see, for example, Persson *et al*. 2005). The emitted laser shot interacts with the surface, generating the backscatter, and the received signal is recorded as a function of time. This signal can contain one or more peaks, which correspond to distinct reflections of the laser beam. In LiDAR applications, the term 'signal intensity' refers to the return amplitude or energy of one echo.

The benefits of using LiDAR intensity have been studied in different application areas, for example in forestry and glaciology, where the signal intensity is already used as additional data source for surface classification and object detection. The fact that the intensity is already delivered to most end-users, the advantages of the active sensor system and its specific wavelength in the NIR generates the potential for using LiDAR intensity (Höfle & Pfeifer 2007). The intensity values are available as attributes for the geometry (x, y, z, I) and in comparison to raw digital images which are typically already georeferenced (Filin 2003; Kager 2004). Ortho-rectified intensity images can be easily produced; they are insensitive to light conditions to a greater extent, e.g. solar irradiation, clouds, illumination shadows and support surface classification where a good spectral separability is given in the NIR (Wolfe & Zissis 1993). Most of the current LiDAR systems record the intensity using small-footprint scanners, which operate with a beam divergence in the range of 0.3–0.8 mrad and a flying altitude above ground of up to 3,500 m. With a beam divergence of 0.8 mrad and a flying height of 1,000 m, the laser footprint diameter would be 0.8 m (and 2.8 m for 3,500 m flying height). For further processing of the intensity in both the

point cloud and in rasterized image formats, the intensity has to be corrected for the influences of topography and flying altitude. The intensity data provided by current commercial systems offer a resolution of 8-bit, 12-bit or 16-bit. There is no detailed insight given into the proprietary pulse detection algorithms. The intensity may correspond to a specific amplitude of the detected echo, e.g. its maximum, but also to the integral of the returned signal over the pulse width. The need for normalized intensity values and images, respectively, is most obviously given for large LiDAR datasets containing strong elevation differences, as for example in high mountainous areas where uncorrected intensity images can hardly be used. Traditionally, intensity information has not always been chosen for feature extraction from LIDAR data because such data can be noisy. Even though the intensity values returned by the scanning unit can be noisy, road material is typically uniform and, as such, it can be detected. By searching for a particular intensity range, it is possible to extract all LiDAR points that were on the road (see, for example, Clode *et al.* 2004).

Höfle & Pfeifer (2007) proposed two independent methods for correcting airborne laser scanning intensities. The first approach is based on the data-driven method and it performs the least-squares adjustment for a given empirical model including intensity and range, the major variable influencing the received signal intensity. The best results were achieved with a model representing a range-square dependency. The second approach uses the radar equation, which describes the loss of emitted pulse power. Both correction methods achieve a significant reduction of local intensity variation within cells of a regular grid (1 m, 3 m and 5 m size) spanned over the study area and an even more significant global adjustment of the single flight strips.

It is now a widely accepted fact that a precise, rational and efficient treatment of geospatial information is essential for practical use. For example, Tsubaki & Fujita (2010) have demonstrated how important is the accuracy of the generated grid with different grid spacing and grid type for the representation of an urban topography. They proposed a method that can generate a fine unstructured grid and incorporate the complicated urban land features precisely without exhausting labour for data preparation. Typically, the processing of LiDAR data involves the application of techniques such as thinning, filtering and interpolation. Thinning (or reduction of data points) is achieved by removing neighbouring points that are found to be within a specified elevation tolerance. The automatic detection and interpretation of bare earth and objects is called a filtering technique. Currently, there are several algorithms available that can be used for filtering purposes. Previous work of Abdullah *et al.* (2011) has evaluated seven algorithms and identified that three of them have a number of issues that make them not suitable for urban flood modelling applications. Out of the remaining four algorithms the Morphological algorithm (i.e. Morph) was singled out as the best for this purpose and it was further improved to incorporate different forms of buildings (buildings with basement, passage buildings and solid buildings) and verified using the data from the Kuala Lumpur case study. Such an improved version of the Morph algorithm has been referred to as the Modified Progressive Morphology Algorithm 1 (MPMA1). Subsequent to that, further improvements of this algorithm were obtained and the new version, which is now referred here as MPMA2, is presented and discussed in the following sections of this paper.

MODIFICATION OF THE PROGRESSIVE MORPHOLOGICAL ALGORITHM 1 (MPMA1)

The work concerning the development of the MPMA1 algorithm has been discussed in detail in Abdullah *et al.* (2011). In the present work, the MPMA1 algorithm has been further improved by focusing on the handling methods for elevated roads and bridges and it has led to a new algorithm, MPMA2.

The overall aspects of the MPMA2 algorithm are:

1. The ability to detect buildings and categorize them into solid buildings, passage buildings and buildings with storage (already incorporated as part of the MPMA1 algorithm).
2. The ability to detect elevated road/train lines, to remove them from a DTM and to incorporate underneath structures.
3. The ability to detect and cross-reference the location of bridges for the purposes of setting up a 1D model and to remove them from the DTM that will be used to set up the 2D model.
4. The ability to apply a data fusion concept and combine the river polygon data with the LiDAR data for the

purpose of identifying riverbanks and interpolating the points between their sides.

A discussion concerning the first aspect, which is the detection of buildings, their classification and representation within a DTM, is given in Abdullah *et al.* (2011). The present paper describes efforts to incorporate the other three aspects. The entire methodology described here was implemented in a Visual Basic code and, as such, embedded in a smaller software application. The flowchart depicting the difference between the previous and the present work is given in Figure 1. The work was carried out in three main steps:

- *Step 1*: Detection and removal of elevated roads and train lines.
- *Step 2*: Incorporation of piles below elevated roads and trains lines.
- *Step 3*: Detection and removal of bridges and the interpolation between the left and right bank of the river (as the river was modelled with the 1D model).

Detection and removal of elevated roads and rail lines

In reference to the detection of buildings, points were labelled as 'High' or 'Low' and 'Steep' or 'Slight' based on

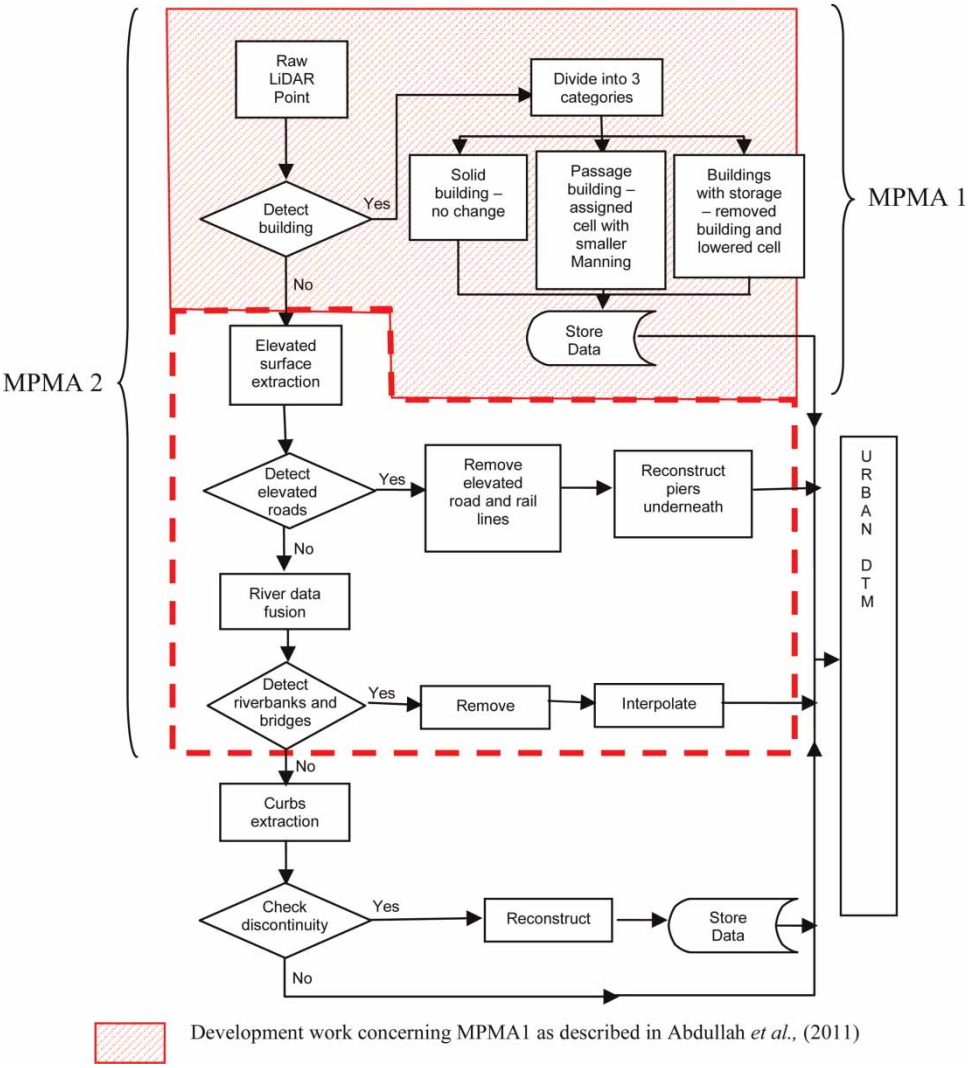

Figure 1 | Flowchart for the development of MPMA1 and MPMA2 algorithms.

their height and associated slope, respectively. The value for 'High' points is made to be user-defined: however, the default value was set to be equal to the maximum value of the average surface (bare earth). Values lower than the value of 'High' were considered as 'Low'. Members of the 'Steep' class are those points which represent vegetation and walls of buildings, while members of the 'Slight' class are representatives of relatively flat surface areas which are based on the following assumption:

- Terrain slopes are usually different from the slopes occurring between the ground and the top of an object (elevated road/train), and
- The elevated road/rail is a highly elevated object in a scene.

The combinations of points with 'High' and 'Slight' labels are points that represent relatively flat areas on high surfaces, including elevated roads, elevated train lines and building roofs, as shown in Figure 2. This concept is used as a basis for detecting elevated roads and trains lines. The next step in this methodological framework is to separate the elevated road/rail lines from other objects. The separation is performed by selecting the points that have intensity values within an acceptable range for the type of road material being detected (in this case asphalt/bitumen), and for this the intensity feature of LiDAR is used. Asphalt is a sticky, black and highly viscous liquid or semi-solid substance that is present in crude petroleum. The primary use of asphalt is in road construction, where it is used as the glue (or binder) for the aggregate particles. It was identified in Lorenzen & Jensen (1989) that 'asphalt' (which in Asia is referred to as bitumen) has the intensity value of 10–20%, 'grass' approximately 50%, 'trees' 30–60% and 'house roofs' 20–30%. As each of these classes has a different intensity value, separation between them can be relatively easily established. As mentioned above, the intensity value of 'asphalt' is approximately 10–20%. By using the typical intensity value for 'asphalt' (i.e. 10–20%), roads were then identified from the points cloud.

In order to see the correlation between intensity values and road objects, 2908 sample points from several areas that represent roads in the points cloud were extracted. These sample points were checked for their intensity values. From this operation it was found that, even though there were some errors, the majority of the sample points (81%) show the values in the range between 10–20%, which corresponds to the range of intensities of asphalt values (Table 1). Figure 3 illustrates the relation between point cloud data and the data that have the intensity values of roads. No values above 30% were recorded as the sample given in Figure 3 concerns only the road area.

Even though the intensity values returned by the scanning unit were noisy, sections of road material were typically uniform for road/rail lines and, as such, they can

Figure 2 | Aerial image of the study area (right) and points with 'High' and 'Slight' labels (with darker colour) which indicate relatively flat areas on high surfaces (left).

Table 1 | Analysis of intensity values

Type	Number of points	Percentage
Intensity (10–20%)	2,369	81.46
Others	539	18.54

be distinguished relatively easily. By searching for a particular intensity range it is possible to extract the points that refer to road/rail lines on the elevated surface.

Incorporation of piles underneath the elevated road/rail lines

The new algorithm allows for incorporation of piles underneath the elevated road/train lines, if the information is available. Within the algorithm, this process is made optional so that if the information about the dimension of piles, the distance between them and their location are not available, the filtering and classification process can still be performed. In this option, points that have been identified as elevated road/train lines are converted into vector form as lines, before they are removed from the DTM. Using these lines as a basis, the pile shapes are incorporated in the vector form, based on locally surveyed information. The information needed for this reconstruction process includes the pile's height, width, length and the distance between the piles. The algorithm also permits the piles to be placed in the middle or perpendicular position with the line, as shown in Figure 4. Once the vector reconstruction process was complete, the created polygons were converted into a grid and then merged back with the DTM.

Removal of bridges across the river and interpolation between river banks

The process of detection and removal of bridges and river surface interpolation has been implemented within the new algorithm by using the data fusion concept. The river polygon data were overlaid over the point cloud data using the buffer of 5 m on both sides of the river and the points within the river polygon were then extracted. From the selected points, the points with intensity values that correspond to asphalt (i.e. 10–20%) were used to identify the location of bridges and to remove them from the DTM. Information about the location of bridges was cross-referenced with the geometry of the 1D model to ensure that all bridges (and culverts) are correctly incorporated within the 1D model. The river banks are then interpolated between the left and right bank elevation values by applying the Kriging interpolation method and the resulting DTM was used in setting up the 2D model and coupling with the 1D model.

CASE STUDY

The study area concerns part of the Klang River basin. It is located on the west coast of Peninsular Malaysia, in the Federal Territory of Kuala Lumpur. The Klang River basin is the most densely populated area in the country, with an

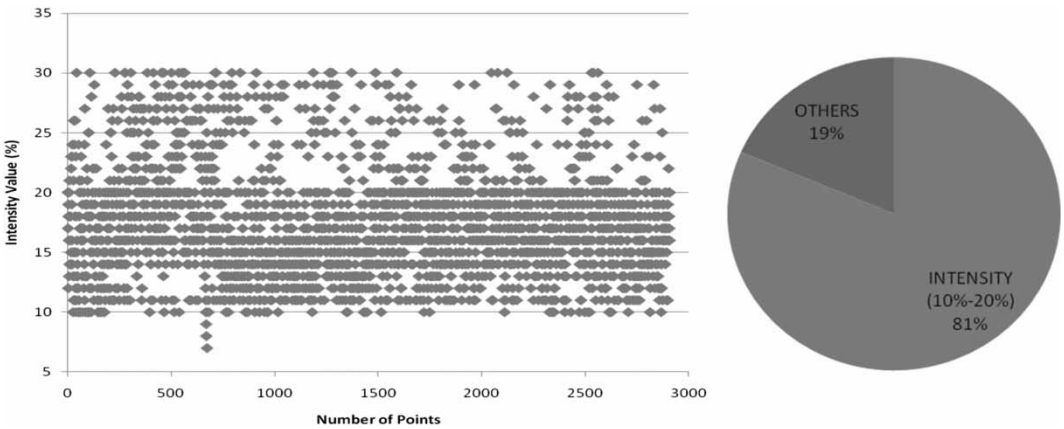

Figure 3 | Illustration of the relation between point cloud data and the data that have intensity values of roads.

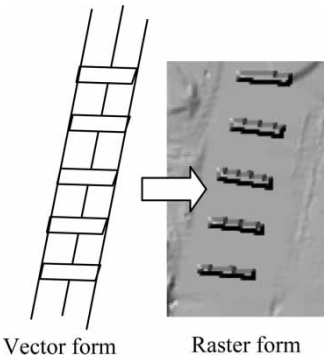

Figure 4 | Reconstruction of piles in vector (left) and raster (right) forms.

estimated population of over 3.6 million, which is growing at almost 5% per year. Although major flood relief works within Kuala Lumpur have been implemented, flooding in the city is still frequent and severe. This causes disruption to various activities in the city, as well as extensive losses suffered by the urban population. The Government of Malaysia, through the Department of Irrigation and Drainage (DID), has commissioned several flood mitigation and river environment enhancement programmes. Recently, a field survey has been carried out in the study area in order to gather information about the piles beneath elevated roads and to use such information for incorporation in a DTM. Survey data concerning typical pile geometry in the study area can be summarized in the following:

- Height, $h = 6$ m
- Length, $l = 2$ m
- Width, $w = 6$ m
- Distance between piles, $lb = 32$ m.

Comparison of urban DTMs

In relation to the treatment of elevated roads/train lines, the comparison of DTMs produced by MPMA1 and MPMA2 algorithms was undertaken at two typical locations (Figure 5).

Location 1 in Figure 5 is situated in the north-west of the study area and has many elevated roads. Location 2 is situated at the downstream end of the study area and contains a typical bridge structure. Comparison of DTMs produced by the two algorithms is given in Figure 6 and discussed in the following section of the paper.

Evaluation of algorithms

A qualitative assessment was undertaken to evaluate MPMA1 and MPMA2. In this assessment, criteria were used that focus on the removal of elevated roads and bridges. The assessment was done by visually analyzing the performance and giving a mark with a weighted value. This weighted value is based on the filter performance in removing the features. If more than 75% of the features are removed the filter is given one mark; if

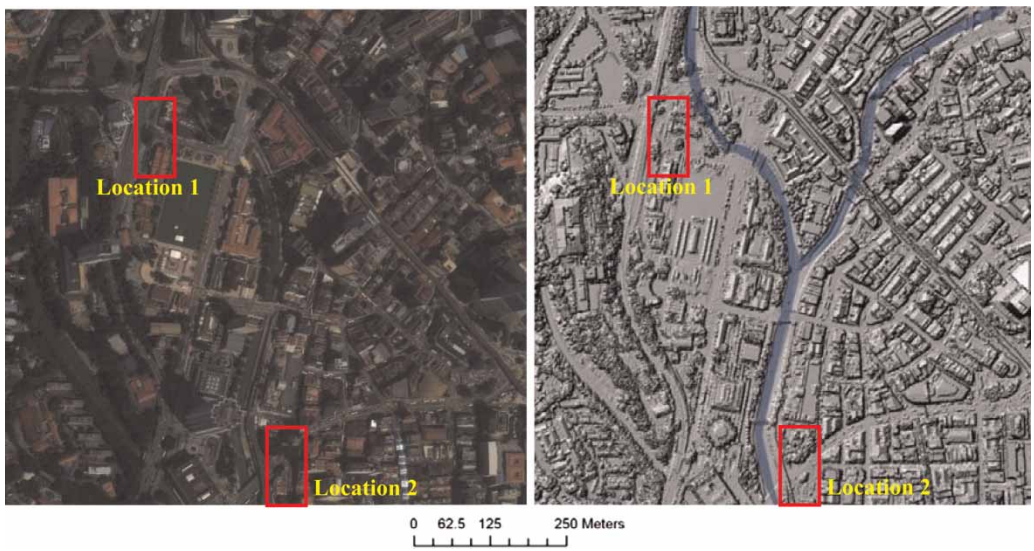

Figure 5 | Locations for comparison of DTMs produced by MPMA1 and MPMA2 algorithms.

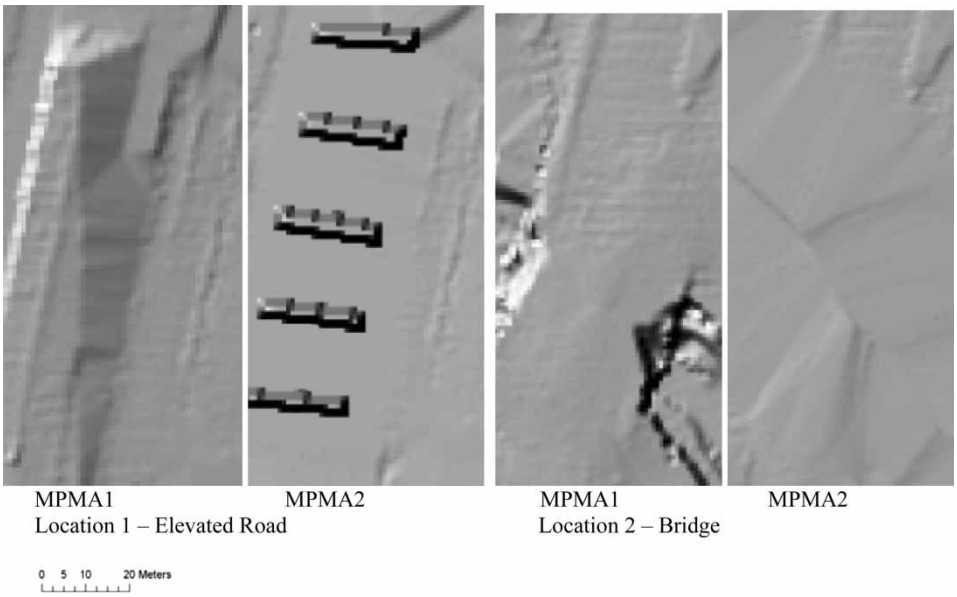

Figure 6 | Comparison of DTM details produced by two algorithms at two locations.

50–75% of the features are removed the filter is given two marks; if 25–50% of the features are removed the filter is given three marks and if less than 25% of the features are removed the filter is given four marks. The least total mark gives an indication which filter performs best with respect to the selected criteria. The DTMs were compared in terms of their efficiency in removing elevated roads and bridges in two locations (Figure 6). Further to the improvements described above, the DTM generated by the MPMA2 algorithm proved to have better results than its predecessor, MPMA1. The DTM generated by MPMA2 appeared as almost flat in all two locations and it managed to incorporate piles in location 2. MPMA1 could remove only 40–70% of the features showing that the two filters act differently and have different capabilities. From the overall analysis of results given in Table 2 it can be noted that MPMA2 produced better results than MPMA1. This can also be followed by observing Figure 6, which illustrates the quality of DTMs obtained by the two algorithms at the two locations (one at the location with the elevated road and the other at the location with the bridge structure).

Table 2 | Comparison of performance of the two algorithms

Indicator	MPMA 1	MPMA 2
Bridge removal	2	1
Elevated road removal	3	1
River alignment	2	1
Building handling	1	1
Total weighted value	8	4

Weighted value: 1 = Excellent, 2 = Fair, 3 = Acceptable, 4 = Poor.

Modelling framework

The model of the study area contains the river network and urban floodplains for which the 1D/2D commercial software packages MIKE 11/MIKE 21 (i.e. MIKEFLOOD) were utilized. The MIKE 11 model of the river used in the present work has been developed and calibrated in a previous study carried out by DHI Water & Environment (2004). The main purpose of that study was to investigate different flood mitigation scenarios. The study also involved a flow/rainfall measuring campaign during the period 1 March 1999–1 April 2003. The time series data used for calibration purposes included nine rainfall events and the results obtained were assessed to be adequate. The details of this work are described in the study report (DHI Water & Environment 2004).

The drainage channel network and the river (including bridges and culverts) were modelled with a 1D model

(MIKE 11). The calculated subcatchment discharges are introduced as lateral or concentrated inflows into the branches of the 1D model network. The hydrographs generated for each subcatchment were calculated using the Nedbor-Afstroming Model (NAM) rainfall–runoff model and nonlinear reservoir method. NAM is a lumped, conceptual rainfall–runoff model used for simulating overland flow, interflow and base flow as a function of the water storage in these mutually interrelated storages representing the storage capacity of the catchment. The floodplain flows were modelled with 2D (MIKE 21) models which were built using different DTMs generated by the MPMA1 and MPMA2 algorithms with spatial resolution of 1 m and coupled with a 1D MIKE 11 model. Both 2D models have the same distribution of Manning's roughness coefficients across the modelled domain. The determination of Manning values has been discussed in detail in Abdullah *et al.* (2011).

Further to the discussion given in the previous section, the 2D model built from the MPMA2 algorithm has managed to remove elevated roads, whereas the MPMA1 algorithm has not been able to do so. Furthermore, the 2D model built from the MPMA2 algorithm also contains 'solid blocks' (i.e. cells with elevations of 6 m higher than the terrain elevation) in places where the piles under elevated roads are located, whereas the 2D model built from the MPMA1 algorithm does not have such a representation (see Figure 6, location 1). In both of the 2D models the buildings were incorporated in an equal way (buildings with basement, passage buildings and solid buildings) following the previous work (see Abdullah *et al.* (2011)). The two sets of 1D/2D models were finally obtained:

- The MIKE 11 model of the drainage channel network and the river (including bridges and culverts) and the MIKE 21 model representing the floodplain built using the MPMA1 algorithm, and
- The MIKE 11 model of the drainage channel network and the river (including bridges and culverts) and the MIKE 21 model representing the floodplain built using the MPMA2 algorithm.

In terms of the coupling with the 1D model, the 2D models were set to the bank-full level of the 1D model. Upstream and downstream boundary conditions of the river model were derived from the results of the previously modelled Klang River model and introduced as inflows (upstream) and water level (downstream) into the model used in the present work.

The rainfall event that occurred on 10 June 2003 was used in model simulations of the present work. The rainfall data of that event was gathered from two rainfall stations: JPS Wilayah and Leboh Pasar. The recorded rainfall at Station JPS Wilayah was 86 mm during a period of 4 h while the Leboh Pasar station recorded 125 mm for the same duration. To account for spatial variability of the rainfall the subcatchments of the 1D model were divided into two groups according to their location and proximity to the two rain-gauge stations and, as such, they were assigned the respective rainfall time series. The model results were then analysed in terms of flood depths and extent of flooded area, and compared against the observations collected at several locations. This observation data was obtained from the Drainage and Irrigation Department of Malaysia (DID).

Discussion of model results

The model results were analyzed in terms of flood depths and extent of flooded area and compared against observations taken at eight locations along the streets, shown in Figure 7. Table 3 indicates the difference between the measured and modelled data at these locations. At six out of eight locations the differences between the measurements and results from the 2D model built from MPMA2 were found to be between 0.2 and 19.1%. In terms of the differences between the measurements and results from the 2D model built from the MPMA1 algorithm, this was found to be in the range between 0.4–24.7%. However, the differences between the measurements and results from the 2D model built from MPMA2 at the locations Dataran Merdeka (1) and Jalan Parlimen (5) were found to be 74.2% and 58.6%, respectively. For the 2D model built from the MPMA1 algorithm, the differences at the same locations were 84.6% and 71%. The reason for the more significant difference for both models can be largely attributed to the effects of the spatial variability of rainfall, lack of high resolution rainfall data and proximity of measurements taken at locations 1 and 5. This was certainly not the case with the other measurement locations as they were situated in close proximity to one of the two rain-gauge stations.

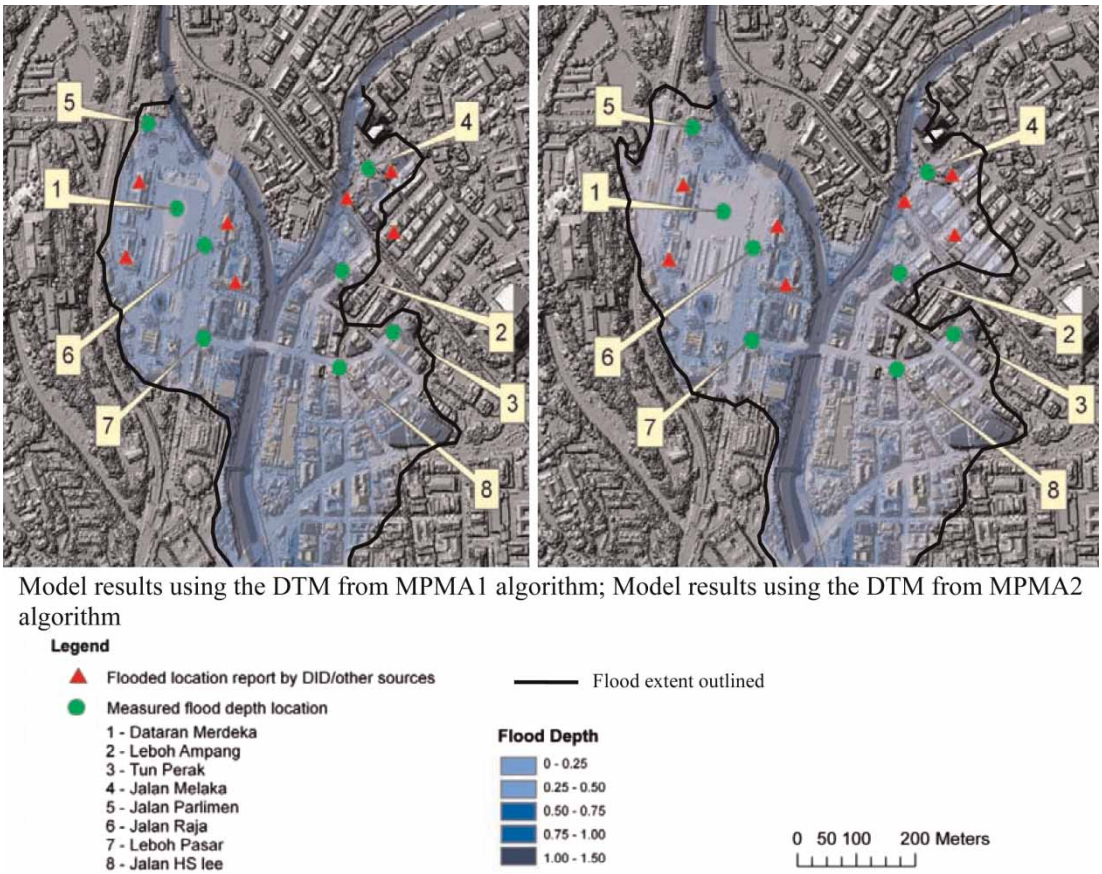

Figure 7 | Modelled and observed flood locations for 10th June 2003 rainfall event. Images illustrate predictions by models with DTMs generated from MPMA1 (left) and MPMA2 (right) algorithms. Observed locations are represented with triangles and circles.

Table 3 | Summary of modelled and measured flood depths for 10 June 2003 rainfall event. The error values were calculated with the following formula: [ABS(Model-Measured)/Measured]*100

Flood depth (m)	Location 1 Dataran Merdeka (1)	% Diff.	Leboh Ampang (2)	% Diff.	Tun Perak (3)	% Diff.	Jalan Melaka (4)	% Diff.
Measured	0.50		1.20		1.00		1.30	
1D/2D model with MPMA2 DTM	0.87	74.2	1.20	0.0	0.98	1.6	1.55	19.1
1D/2D model with MPMA1 DTM	0.92	84.60	1.19	0.7	0.96	3.6	1.62	24.7
Flood depth (m)	**Location 2 Jalan Parlimen (5)**	**% Diff.**	**Jalan Raja (6)**	**% Diff.**	**Leboh Pasar (7)**	**% Diff.**	**Jalan HS Lee (8)**	**% Diff.**
Measured	0.50		1.00		0.65		1.00	
1D/2D model with MPMA2 DTM	0.79	58.6	0.93	6.7	0.73	12.9	0.98	2.0
1D/2D model with MPMA1 DTM	0.85	71.0	0.88	12.3	0.75	15.1	0.99	0.4

The flood extent shown in Figure 7 is due to the combined effects of river-related overbank discharges, as well as discharges from the inland drainage system. From Figure 7 it can be observed that the model results using the DTM from the MPMA1 algorithm have generated flood areas with larger depths and smaller extents when compared to

the model results using the DTM from the MPMA2 algorithm. This difference can be explained by the fact that the results obtained from the model that uses the DTM from the MPMA1 algorithm have more local obstructions due to the algorithm's inability to remove elevated roads.

For the overall study area (1 km^2), the results from the 2D model built from the MPMA2 algorithm show that 45% of that area (or 0.45 km^2) is flooded, whereas the results from the 2D model built from the MPMA1 algorithm show that 30% of that area (or 0.3 km^2) is flooded. In terms of the more localized comparison, areas covered in locations A (0.07 km^2), B (0.07 km^2) and C (0.07 km^2) show that the results from the 2D model built from the MPMA2 algorithm have predicted a flood extent of 45% (or 0.03 km^2), 55% (or 0.04 km^2) and 52% (or 0.03 km^2), respectively (see Figures 8 and 9). In terms of the results obtained from the 2D model built from the MPMA1 algorithm, the flooded areas at locations A, B and C are 30% (0.02 km^2), 48% (0.03 km^2) and 35% (0.02 km^2), respectively (Figures 8 and 9). Table 4 gives a summary of these results.

At all three locations, it can be noted that the 2D model built from the MPMA2 algorithm has generated a wider flood extent, which is due to the removal of elevated roads and bridges in the surrounding area, thus allowing the water to flow through. Additionally, the wider extent computed by the 2D model built from the MPMA2 algorithm in location B is due to its geographical position located at the most downstream part of the study area. This location is not only influenced by the removal of the features in close proximity but also by the cumulative effects from the entire area.

From the analysis of computed flood velocities, it can be observed that the results from the 1D/2D model that uses the DTM from the MPMA2 algorithm correspond better to reality as the DTM better represents the terrain characteristics (Figure 10). In the 1D/2D model that uses the DTM from the MPMA1 algorithm it is obvious that the water cannot flow through the cells where the elevated roads are located as they represent obstructions and create diversions. This is also evident from observations of velocity vectors given in Figure 10. The difference in results indicates that raw LiDAR data need to be carefully processed, otherwise inappropriate representation of features in a DTM can generate misleading results.

CONCLUSIONS

The paper describes efforts to improve one of the LiDAR filtering algorithms with the main objective to enable more accurate 1D/2D urban flood modelling. The algorithm selected for improvement is the MPMA1 algorithm which

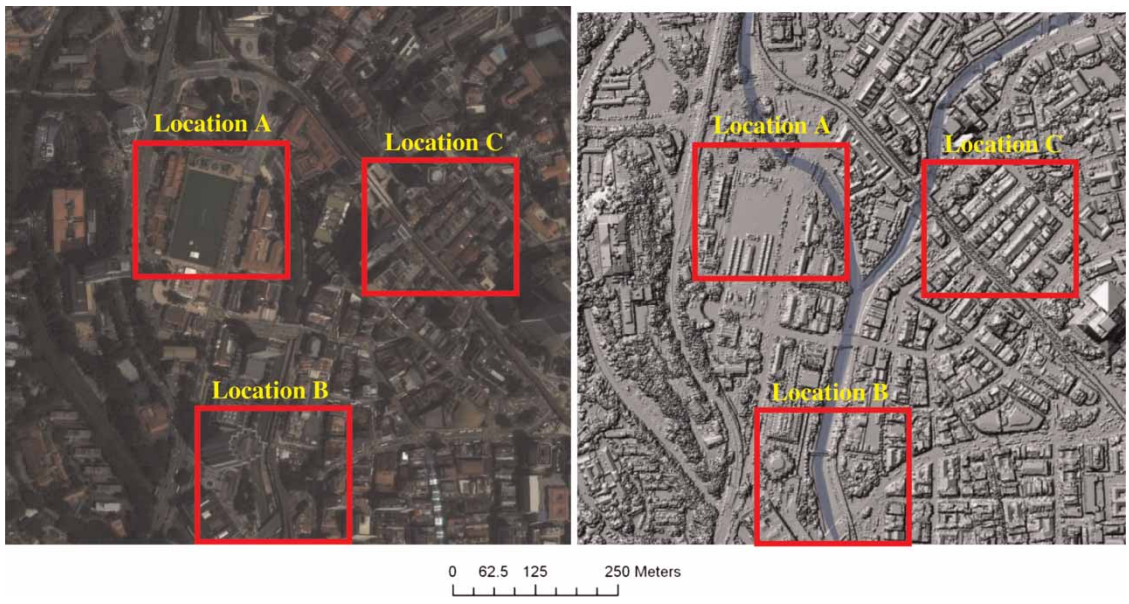

Figure 8 | Locations where flood extents were compared.

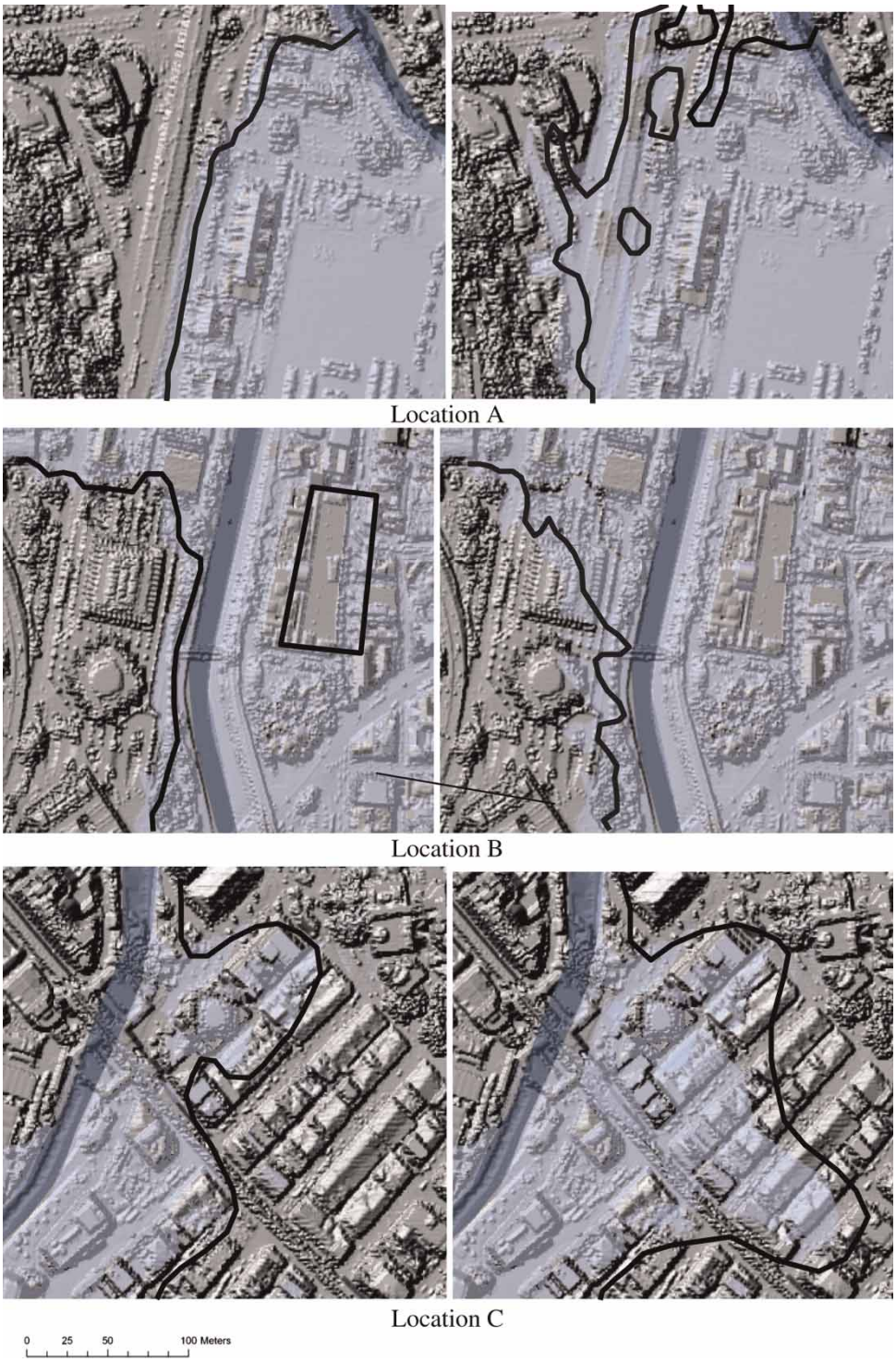

Figure 9 | The difference between flood extents produced by the 2D models built from the MPMA1 algorithm (left) and MPMA2 algorithm (right). The flood extents have been superimposed onto the original DSM of study area using the 'hillshade' function.

Table 4 | Summary of modelled flood extents for 10 June 2003 rainfall event

	Overall Area (km^2)	Percent (%)	Location A Area (km^2)	Percent (%)	Location B Area (km^2)	Percent (%)	Location C Area (km^2)	Percent (%)
2D model result from MPMA1	0.30	30	0.02	30	0.03	48	0.02	35
2D model result from MPMA2	0.45	45	0.03	45	0.04	55	0.03	52
Difference	0.15	15	0.01	15	0.01	7	0.01	17

has already been discussed by the authors in a previous paper and which was originally developed to deal with different kinds of buildings (buildings with basement, passage buildings and solid buildings). That algorithm was further improved to deal with objects such as elevated road/train lines and bridges. Such geometric 'discontinuities' can play a significant role in diverting the shallow flows that are generated along roads, through fences and around buildings. The present work has led to the development of an improved version of the MPMA1 algorithm which has been referred to as the MPMA2 algorithm. The new algorithm has been implemented in a Visual Basic code and embedded in a software application. The key aspects of this algorithm are: the ability to deal with different kinds of buildings, the ability to detect elevated road/train lines and represent them in accordance with the actual reality and the ability to deal with bridges and riverbanks. The first aspect has been dealt with in the previous work concerning the development of the MPMA1 algorithm, whereas the other aspects are original features of the MPMA2 algorithm.

In terms of the detection of elevated road/train lines, the data fusion method was combined with the analysis of LiDAR intensities and those points that fall within the range of asphalt intensities were identified. This was done by using intensity, height, and slope information derived from the LiDAR data. Automatic procedures were developed within the code to remove the elevated roads/train lines and to incorporate the underneath piles. Furthermore, detection of bridges and their cross-referencing against geometric details of the 1D model was also undertaken. After that, the bridges were removed from the DTM (as the river was modelled with the 1D model), and the left and right banks of the river were interpolated using the Kriging method.

The modelling framework applied in the present work involved building the bathymetry for 2D models using two filtering algorithms (MPMA1 and MPMA2), setting up of 2D models within the MIKE 21 system and coupling them with an existing 1D model (MIKE 11) which was developed in a previous study of Kuala Lumpur. Following this, the two sets of 1D/2D models were obtained and simulated for a rainfall event that occurred on 10 June 2003. The model results were then analysed in terms of flood depths and extent of flooded area, and compared against the observations collected at several locations.

The overall comparison of results suggests that the results of a 2D model built from the MPMA2 algorithm are in a closer agreement with the measurements than the results of a 2D model built from the MPMA1 algorithm. The difference in flood extents produced by two models was found to be of the order of 15% of the total area. In terms of the flood depth comparisons, at six out of eight locations the differences between the measurements and results from the 2D model built from MPMA2 were found to be in the range between 0.2–19.1%. In terms of the differences between the measurements and results from the 2D model built from the MPMA1 algorithm, this was between 0.4–24.7%. However, at two locations, these differences were found to be more significant for both models and they were attributed to the effects of spatial variability of rainfall, lack of high resolution rainfall data and proximity of measurements taken at identified locations.

The paper demonstrates that the terrain topography and incorporation of urban features represent a factor that could make for substantial differences between the results obtained from models with different level of details and the corresponding DTMs. This highlights the need for careful processing of raw LiDAR data and cautious generation of a DTM which is aimed to be used in urban flood modelling. Furthermore, the requirements for description of terrain features and DTM's resolution must involve careful consideration of the geometrical features of the area under study and the objectives of the study.

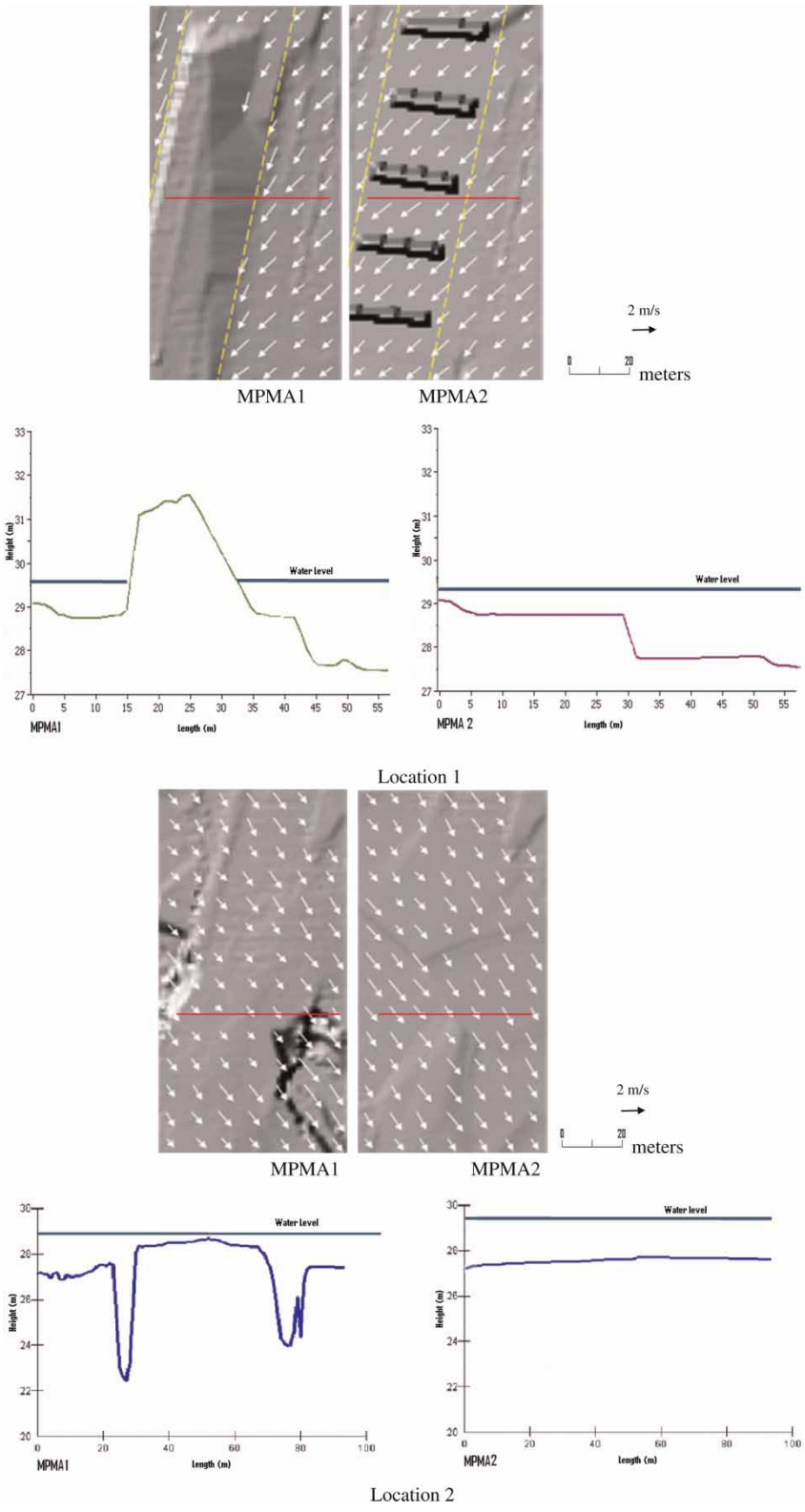

Figure 10 | Comparison of velocity vectors and DTMs generated by MPMA1 and MPMA2 algorithms and their respective cross sections taken at locations 1 and 2.

REFERENCES

Abdullah, A. F., Vojinovic, Z., Price, R. K. & Aziz, N. A. A. 2011 A methodology for processing raw LiDAR data to support urban flood modelling framework. *J. Hydroinf.* (in press).

Chen, A. S., Hsu, M. H., Teng, W. H., Huang, C. J., Yeh, S. H. & Lien, W. Y. 2005 Establishing the database of inundation potential in Taiwan. *Nat. Hazards* **37** (1–2), 107–132.

Clode, S., Kootsookos, P. & Rottensteiner, F. 2004 The automatic extraction of roads from Lidar data. In *ISPRS, Istanbul, Turkey, 20th Congress, Commission 3*. Elsevier, Amsterdam, pp. 231–236.

DHI Water & Environment (M) Sdn. Bhd. 2004 *Klang River Basin Environment Improvement and Flood Mitigation Project (Stormwater Management and Road Tunnel – SMART)*. Final Report to Government of Malaysia.

Djordjevic, S., Prodanovic, D., Maksimovic, C., Ivetic, M. & Savic, D. A. 2005 SIPSON – simulation of interaction between pipe flow and surface overland flow in networks. *Water Sci. Technol.* **52** (5), 275–283.

Evans, B. 2008 Automated bridge detection in DEMs via LiDAR data sources for urban flood modelling. In *Proc 11th Int. Conf. on Urban Drainage, Edinburgh, August*. IWA, London, pp. 1–10.

Filin, S. 2003 Recovery of systematic biases in laser altimetry data using natural surfaces. *Photogramm. Engng. Remote Sensing* **69** (11), 1235–1242.

FLI-MAP 400 Specifications 2010 *John Chance Land Surveys, Inc. Research Centre*. Available from: http://www.flimap.com/site47.php (accessed 27 December 2010).

Garcia-Navarro, P. & Brufau, P. 2006 Numerical methods for the shallow water equations: 2D approach. In: *River Basin Modeling for Flood Risk Mitigation* (D. W. Knight & A. Y. Shamseldin, eds.). Taylor & Francis, London, pp. 409–428.

Höfle, B. & Pfeifer, N. 2007 Correction of laser scanning intensity data: data and model-driven approaches. *ISPRS J. Photogramm. Remote Sensing* **62**, 415–433.

Houzelle, S. & Giraudon, G. 1992 *Automatic Feature Extraction and Localization using Data Fusion of Radar and Infrared Images*. AGARD, Radio Location Techniques. NASA Astrophysics Data System.

Hunter, N. M., Bates, P. D., Horritt, M. S. & Wilson, M. D. 2007 Simplified spatially-distributed models for predicting flood inundation: a review. *Geomorphology* **90**, 208–225.

Hunter, N. M., Bates, P. D., Neelz, S., Pender, G., Villanueva, I., Wright, N. G., Liang, D., Falconer, R. A., Lin, B., Walle, S., Crossley, A. J. & Mason, D. C. 2008 Benchmarking 2D hydraulic models for urban flooding. *Wat. Mngmnt.* **161** (1), 13–30.

Kager, H. 2004 Discrepancies between overlapping laser scanning strips – simultaneous fitting of aerial laser scanner strips. *Int. Arch. Photogramm. Remote Sensing Spatial Inf. Sci.* **35** (Part B1), 555–560.

Kuiry, S. N., Sen, D. & Bates, P. D. 2010 A coupled 1D-quasi 2D flood inundation model with unstructured grids. *J. Hydraul. Engng.* **136** (8), 493–506.

Lorenzen, B. & Jensen, A. 1989 Changes in leaf spectral properties induced in barley by cereal powdery mildew. *Remote Sens. Environ.* **27**, 201–209.

Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, S. B. & Djordjevic, S. 2004 Potential and limitations of 1D modelling of urban flooding. *J. Hydrol.* **299** (3–4), 284–299.

Persson, Å., Söderman, U., Töpel, J. & Ahlberg, S. 2005 Visualization and analysis of full-waveform airborne laser scanner data. *Int. Arch. Photogramm. Remote Sensing Spatial Inf. Sci.* **36** (Part 3/W19), 103–108.

Price, R. K. & Vojinovic, Z. 2011 *Urban Hydroinformatics: Data, Model and Decision Support for Integrated Urban Water Management*. IWA, London.

Schubert, J. E., Sanders, B. F., Smith, M. J. & Wright, N. G. 2008 Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding. *Adv. Water Res.* **31**, 1603–1621.

Sithole & Vosselman 2006 Bridge detection in airborne laser scanner data. *ISPRS J. Photogramm. Remote Sensing* **61**, 33–46.

Trias-Sanz, R. & Lomenie, N. 2003 Automatic bridge detection in high-resolution satellite images. In *Proc. 3rd International Conference on Computer Vision Systems (ICVS 03), Graz, Austria, 1–3 April*. Springer, Berlin, pp. 172–181.

Tsubaki, R. & Fujita, I. 2010 Unstructured grid generation using LiDAR data for urban flood inundation modeling. *Hydrol. Process.* **24**, 1404–1420.

Vojinovic, Z. & Tutulic, D. 2009 On the use of 1D and coupled 1D–2D modelling approaches for assessment of flood damages in urban areas. *Urban Wat. J.* **6** (3), 183–199.

Vojinovic, Z., Seyoum, S. D., Mwalwaka, J. M. & Price, R. 2011 Effects of model schematisation, geometry and parameter values on urban flood modelling. *Water Sci. Technol.* **63** (3), 462–467.

Wang, Y. & Zheng, Q. 1998 Recognition of roads and bridges in SAR images. *Pattern Recogn.* **31** (7), 953–962.

Wolfe, W. L. & Zissis, G. J. 1993 *The Infrared Handbook*. IRIA Series in Infrared and Electro Optics, Environmental Research Institute of Michigan, Ann Arbor, MI.

First received 23 January 2011; accepted in revised form 24 May 2011. Available online 23 July 2011