CrossMark

# Overcoming data scarcity in flood hazard assessment using remote sensing and artificial neural network

Dong-Eon Kim[1,2]*, Philippe Gourbesville[2] and Shie-Yui Liong[1]

\* Correspondence: tmskde@nus.edu.sg
[1]Tropical Marine Science Institute, National University of Singapore, 18 Kent Ridge Road, Singapore 119227, Singapore
[2]Polytech Lab, Polytech Nice Sophia, University of Nice Sophia Antipolis, 930 Route des Colles, 06903 Sophia Antipolis, France

## Abstract

Many urban cities in Southeast Asia are vulnerable to climate change. However, these cities are unable to take effective countermeasures to address vulnerabilities and adaptation due to insufficient data for flood analysis. Two important inputs required in flood analysis are high accuracy Digital Elevation Model (DEM), and long term rainfall record. This paper presents an innovative and cost-effective flood hazard assessment using remote sensing technology and Artificial Neural Network (ANN) to overcome such lack of data. Shuttle Radar Topography Mission (SRTM) and multispectral imagery of Sentinel-2 are used to derive a high-accuracy DEM using ANN. The improvement of SRTM's DEM is significant with a 42.3% of reduction on Root Mean Square Error (RMSE) which allows the flood modelling to proceed with confidence. The Intensity Duration Frequency (IDF) curves that were constructed from precipitation outputs from a Regional Climate Model (RCM) Weather Research and Forecasting (WRF) were used in this study. Design storms, calculated from these IDF curves with different return periods were then applied to numerical flood simulations to identify flood prone areas. The approach is demonstrated in a flood hazard study in Kendal Regency, Indonesia. Flood map scenarios were generated using improved SRTM and design storms of 10-, 50- and 100-year re-turn periods were constructed using the MIKE 21 hydrodynamic model. This novel approach is innovative and cost-effective for flood hazard assessment using remote sensing and ANN to overcome lack of data. The results are useful for policy makers to understand the flood issues and to proceed flood mitigation adaptation/measures in addressing the impacts of climate change.

**Keywords:** Climate model, MIKE 21, Sentinel-2 multispectral imagery, Topography
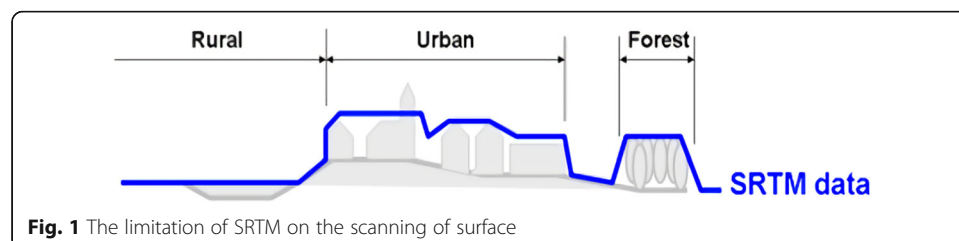
## Introduction

Flood modelling is a useful tool for simulating/predicting water-related situations/disasters. In order to achieve a robust and reliable outcome of flood information, good quality data are required, such as a Digital Elevation Model (DEM) and hydrological/hydraulic data. Among those, DEM is crucial in modelling which reflects the actual topographic characteristics of the catchment (Kim et al. 2018). Bathrellos et al. (2016, 2017) conducted research on major factors affecting natural disaster including urban floods using topographic information such as slope, elevation and distance from streams. This quantitative geomorphological analysis was

able to verify the past flood events well and this implies that DEM is a critical factor in flood assessment. High resolution and high accuracy of DEM has considerable influence in 2D flood modelling (Abily et al. 2015; Ngoc and Gourbesville 2016). Abily et al. (2016) conducted the sensitivity analysis of high resolution topography data in 2D flood modelling and emphasized that the water depth varies up to 1 m based on different resolutions in the study area. There are several sources to download the free DEMs such as the ASTER (Advanced Space-borne Thermal Emission and Reflection Radiometer), SRTM, and GTOPO30 (Global 30 Arc-Second Elevation). The SRTM at 30 m resolution was chosen for this study to derive finer resolution (20 m) data with improved DEM using the developed technique.

The SRTM is an international joint project to collect three-dimensional digital mapping of over 80% of the Earth's surface (between 60° N and 56° S) (USGS 2000). The performance requirements for the SRTM data are such that the linear vertical absolute height error shall be less than 16 m and the relative height error shall be less than 10 m, for 90% of the data (Rodriguez et al. 2005). The SRTM data is free and accessible to the public, however, its accuracy is limited with Root Mean Square Error (RMSE) of approximately 14 m over Singapore's forest areas due to C-band wavelengths ($\lambda \approx 5.6$ cm) that does not adequately penetrate the vegetation canopy (Wendi et al. 2016). Thus, the elevation in vegetation area presents an intermediate height between top of canopy and the bare surface. Also, due to its coarse resolution (30 m), it does not allow to present precise urban characteristics. Therefore, in the areas where high-accuracy surveyed DEM is not available, additional studies are required to obtain reliable topography and to overcome the uncertainty introduced by the measurement protocol. Figure 1 shows the limitation of SRTM on the ground surface scanning (Coude 2018).

Unavailability of high-accuracy DEM and precipitation data are the major limitations of flood hazard assessment in many developing countries. The design of optimal mitigation measures including efficient stormwater systems request long records of hydrological data in order to generate the relevant rainfall IDF curves and to identify the high-risk areas (Liew et al. 2017). Kendal Regency, Indonesia, facing aforementioned data scarcity, is regularly facing flood issues and warrants a risk analysis. The best option for this analysis is to mobilize a 2D hydraulic model (e.g. MIKE 21) in order to simulate the flood propagation. As the study area suffers from lack of rainfall data, the methodology developed by Liew et al. (2012), which used climate model projected rainfall, has been implemented with rainfall data extracted from a high resolution regional climate model, Weather Research and Forecasting (WRF) driven by the European Reanalysis ERA-40 dataset, referred to in this paper as 'WRF/ERA40'.

In this research, flood hazard assessment model was developed to construct the flood maps for Kendal in Indonesia. The publicly available SRTM30 DEM was improved



**Fig. 1** The limitation of SRTM on the scanning of surface

Kim *et al. Smart Water*      (2019) 4:2

Page 3 of 15

using remote sensing imagery incorporated with machine learning technique. The Regional Frequency Analysis (RFA) using the WRF/ERA40 as proxy data method was applied in the study area to generate the IDF curves.

## Methodology

### General concept

The Artificial Neural Network (ANN) is a useful tool to apply pattern recognition (classification) in engineering and scientific applications such as biology, medical science and remote sensing technology (Nguyen et al. 2018). This study utilizes the strength of pattern recognition in ANN to correct/improve the elevation in SRTM. As stated above, the limitation of scanning on the surface due to coarse resolution and penetration of canopy issue, SRTM has a tendency to overestimate or underestimate the elevation in forest and urban areas. The multispectral imagery supports to classify the different types of areas by its reflectance. The ANN is then trained using SRTM and multispectral imagery, together with the target data (which is the high accuracy DEM) to calculate the weights in different type of areas. Once the results are satisfied in the validation part, the trained ANN can be applied to other areas. The generated DEM through ANN, is then used as an input in the flood model with rainfall information derived from RCM output.

### Study area

Kendal regency is under Central Java and has borders with Java Sea in the north, Semarang City in the east, Temanggung in the south and Batang Regency in the west. The northern region of the city is dominated by coasts, while in the south, lined by mountains. The study area is classified as Tropical Rainforest 'Af' based on Koppen-Geiger climate classification (Kottek et al. 2006). The average annual temperature is 20.2 °C and the yearly average rainfall is 3276 mm. The rainfall is lowest in August and highest in January.

The study area is divided into two:

- lowland areas in the northern coast;
- highland areas in the southern mountains.

The elevation ranges from 0 m to 2604 m in SRTM. The location of Kendal regency and its terrain map are shown in Fig. 2.

### Data preparation

In order to utilize the ANN, different types of remote sensing data are required. In this study, TanDEM-X (Wessel 2016), SRTM, surveyed high accuracy DEM and multispectral imagery from Sentinel 2 were used. As the SRTM is attributed with 30 m horizontal resolution in contrast to TanDEM-X (12 m), surveyed DEM (5 m) and Sentinel 2 (10–60 m), all input layers are standardized to 20 m resolution through sampling method.

For the Regional Frequency Analysis (RFA), proxy data from downscaled RCM (30 × 30 km over the study domain) (WRF/ERA40) were used.

**Fig. 2** The location of study area and its terrain map
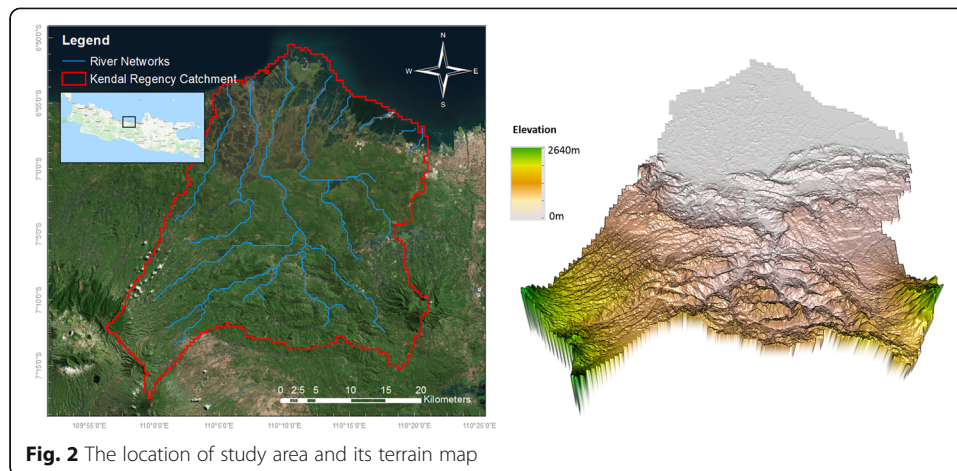
### SRTM Data

The United States (US) National Aeronautics and Space Administration (NASA) released the SRTM data set for the globe (between 60° N and 56° S) at 3 arc-second resolution and 1 arc-second resolution for the US in 2003 (Jarvis et al. 2004). One arc-second resolution for the globe is available after 2015. The SRTM DEM is referenced to mean sea level with the Earth Gravitational Model 1996 (EGM 1996) geoid model. Hence, this study aimed at a higher resolution than its original, resampled to 20 m.

### TanDEM-X Data

The German Aerospace Center (DLR) has been operating Germany's first two formations flying Synthetic Aperture Radar (SAR) satellites, TerraSAR-X and TanDEM-X, with the objective to generate an updated global DEM. The DEM has a spatial resolution of 0.4 arc-second ($\approx$ 12 m) with 2–4 m in relative vertical accuracy. The vertical datum of TanDEM-X is WGS84-G1150 ellipsoidal heights (Wessel 2016) and it has changed to geoid system to standardize the datum system with SRTM. As the DEM has higher accuracy and higher resolution than SRTM, it used for training of ANN after the resampling.

### Sentinel 2 multispectral imagery

Sentinel 2 is an earth observation mission developed by European Space Agency (ESA) to provide the continuity services on multispectral high resolution optical observation over the global terrestrial surface (Drusch et al. 2012). It provides the 13 spectral band spans from the visible (VIS) and the near infrared (NIR) to the short wave infrared (SWIR) at different resolutions ranging from 10 to 60 m with 10 days revisit frequency. Table 1 shows the Sentinel 2 spectral bands definition.

The bands of 2, 3, 4 and 8 are used for land cover classification and 5, 6, 7 and 8A are mainly dedicated to vegetation area. Band 1, 9, 10, 11 and 12 are used for atmospheric correction and cloud screening. In this study, 8 bands (2, 3, 4, 5, 6, 7, 8 and 8A) are used for the ANN nodes with 20 m standardized resolution. Figure 3 shows the different reflectance values based on various land use over Singapore. The reflectance in forest area is higher than urban area at SWIR bands (Band 6–8) while urban area is

**Table 1** Sentinel 2 spectral bands

| Bands | Central wavelength (nm) | Band width (nm) | Pixel resolution (m) |
|---|---|---|---|
| Band 1 – Coastal aerosol | 443 | 20 | 60 |
| Band 2 – Blue | 490 | 65 | 10 |
| Band 3 – Green | 560 | 35 | 10 |
| Band 4 – Red | 665 | 30 | 10 |
| Band 5 – Vegetation Red Edge | 705 | 15 | 20 |
| Band 6 – Vegetation Red Edge | 740 | 15 | 20 |
| Band 7 – Vegetation Red Edge | 783 | 20 | 20 |
| Band 8 – NIR | 842 | 115 | 10 |
| Band 8A – Narrow NIR | 865 | 20 | 20 |
| Band 9 – Water vapour | 945 | 20 | 60 |
| Band 10 – SWIR – Cirrus | 1380 | 20 | 60 |
| Band 11 – SWIR | 1610 | 90 | 20 |
| Band 12 – SWIR | 2190 | 180 | 20 |

higher than forest area at NIR bands (Band 2–5). These different characteristics at each band help to classify land use in ANN as input nodes.

### WRF/ERA40 rainfall grid

The ERA40 re-analyses (Uppala et al. 2005) provide information of climate variables every 6 h, with a horizontal resolution of 2.5° × 2.5° and 23 vertical levels. The data cover more than 40 years (1957 to 2002). Most of the variables are available at a resolution of 2.5° × 2.5° on a regular latitude and longitude grid. The ERA40 reanalysis products use a global spectral grid model and assimilate part of the observational data from a wide variety of observed sources. As rainfall is the most important climate variable and is also a key variable for assessment of future climate changes, model evaluations of simulated rainfall is crucial. However, this is not within the scope of this study
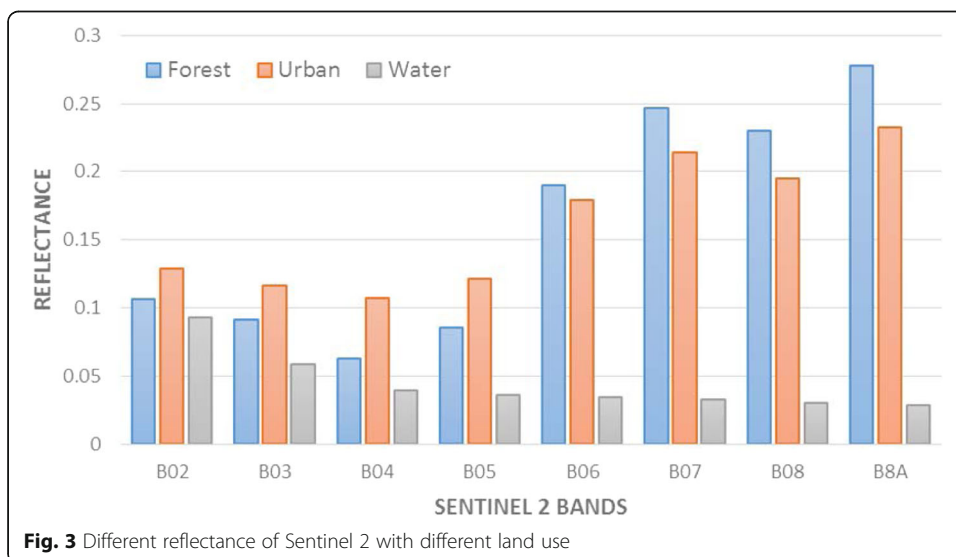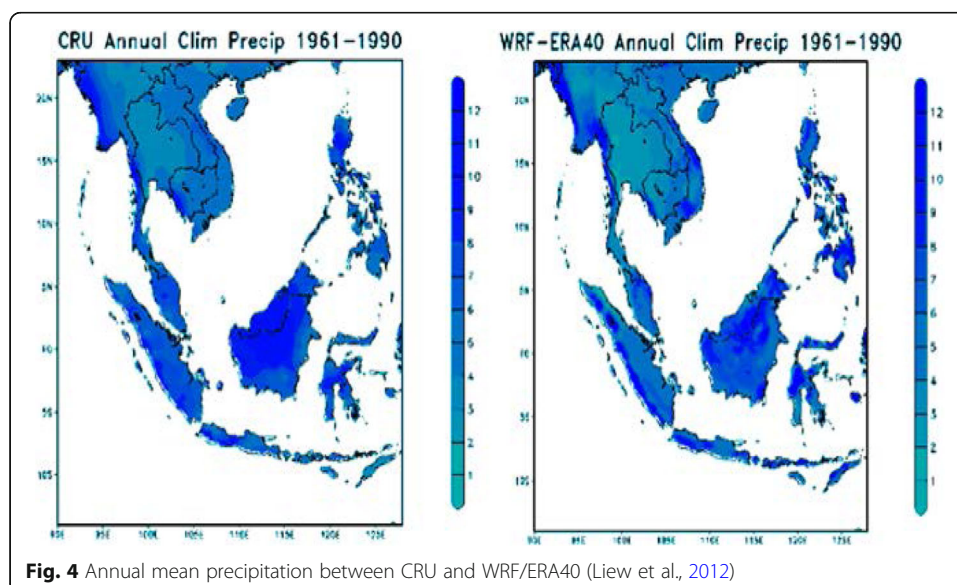


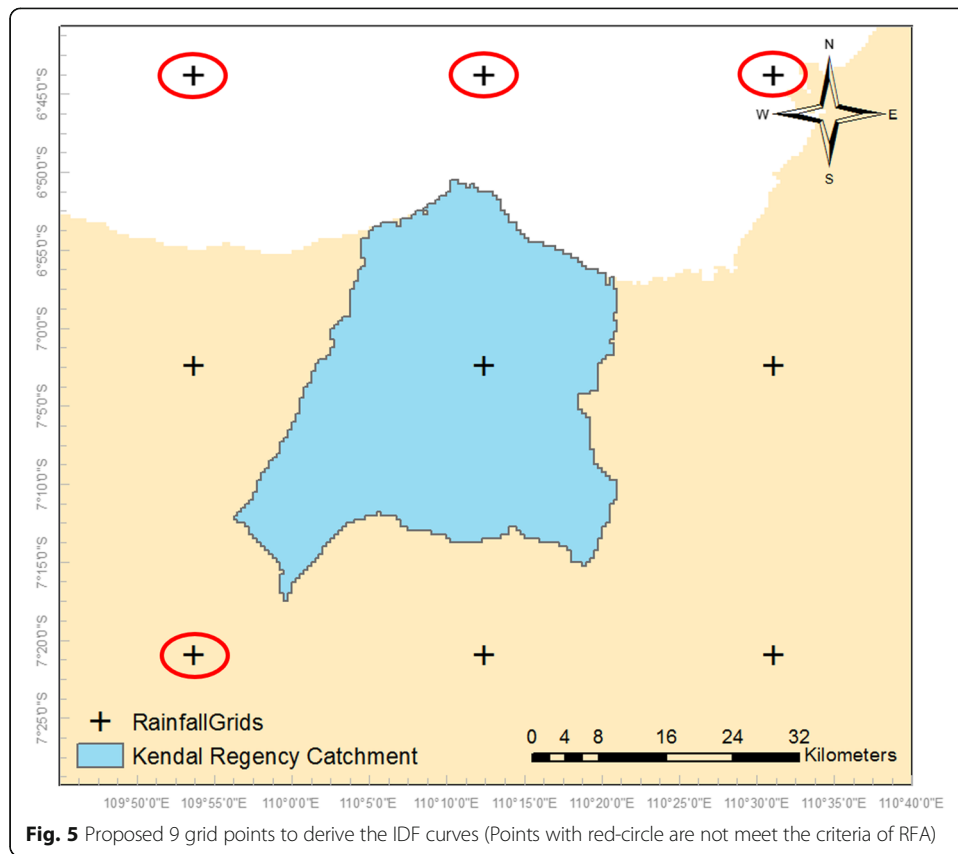**Fig. 3** Different reflectance of Sentinel 2 with different land use

and the reader is referred to Liew et al. (2012, 2014) further. This paper also describes constructing storm design drainage curves from RCM outputs. This study makes use of these IDF curves as the starting point for flood simulations. Nevertheless, for brevity, the model reproduced rainfall is compared against observations, as seen in Fig. 4. After the validation, this study investigates the frequencies and magnitudes of extreme rainfall events from WRF/ERA40 for 9 grid points in the study area (Fig. 5). The baseline period over the present day climate considered was from 1961 to 1990.

### Generation of improved DEM

The procedure of DEM improvement using ANN is divided into three major parts, which are training, validation, and application. ANN is a nonlinear mathematical structure model which is capable of representing an arbitrarily complex nonlinear process that relates the inputs and outputs of any system (Hsu et al. 1995). The ANN mimics the human's brain with two primary features: the ability to 'learn' and 'generalize' from limited information (Hewiston and Crane 1994). Also, ANN is powerful in the analysis of remotely sensed data, particularly in the classification of land use and land cover (Manibhushan et al. 2011).

At the training stage, the weights of ANN are updated by backward propagation when the target/reference data sets are provided. Pontian District in southern Malay Peninsula was selected, in this study, as the training part. SRTM and 8 bands of multispectral imagery from Sentinel 2 were used for input nodes and TanDEM-X was used for target nodes. Table 2 shows the example of ANN input layers. The 8 bands of Sentinel 2 are used to classify the various land uses which can be used to calculate the weights in ANN for different land uses so that the error can be reduced between the elevation of SRTM and TanDEM. The validation was conducted in Singapore where high accuracy DEM is available. The trained ANN was finally applied to the study area in Kendal Regency. The procedure of DEM improvement is shown in Fig. 6.



**Fig. 4** Annual mean precipitation between CRU and WRF/ERA40 (Liew et al., 2012)

**Fig. 5** Proposed 9 grid points to derive the IDF curves (Points with red-circle are not meet the criteria of RFA)

### Generation of IDF curves

The RFA has been conducted using proxy rainfall data from WRF/ERA40. Hosking and Wallis ([1997](#)) developed a complete algorithm for the RFA method based on the approach of L-moments by pooling the sites with similar statistical characteristics in a homogeneous region instead of a single site in the at-site frequency analysis. The discordancy and heterogeneity measures are the primary indicators for accepting or rejecting the grid points from the study area. A discordancy measure is used to identify the sites with gross errors in the data or the data are grossly discordant with the region as a whole (Liew et al. [2014](#)). Also, the heterogeneity is a statistical test to define the regional heterogeneity which is based on L-moments and on the theory that all stations

**Table 2** Input, target and output layers in ANN (example)

| Input layer | | | | | | | | | Target layer | Output layer |
|---|---|---|---|---|---|---|---|---|---|---|
| B02 | B03 | B04 | B05 | B06 | B07 | B08 | B8A | SRTM (m) | TanDEM (m) | Improved DEM (m) |
| 0.0898 | 0.0884 | 0.0604 | 0.0922 | 0.2329 | 0.3059 | 0.2901 | 0.3296 | 26 | 23.95 | To be calculated |
| 0.089 | 0.0865 | 0.0608 | 0.0928 | 0.2408 | 0.3137 | 0.2787 | 0.3416 | 27 | 24.32 | |
| 0.0835 | 0.0749 | 0.0454 | 0.0866 | 0.2321 | 0.2976 | 0.2733 | 0.325 | 27 | 30.24 | |
| 0.0933 | 0.0927 | 0.0715 | 0.0892 | 0.2442 | 0.3112 | 0.2762 | 0.3503 | 27 | 25.55 | |
| 0.0879 | 0.0797 | 0.0517 | 0.0676 | 0.2178 | 0.272 | 0.2666 | 0.308 | 26 | 25.08 | |
| 0.0856 | 0.0761 | 0.0496 | 0.0834 | 0.2172 | 0.2736 | 0.2367 | 0.3096 | 25 | 29.17 | |
| 0.0944 | 0.0901 | 0.0695 | 0.0976 | 0.2332 | 0.2924 | 0.2684 | 0.3247 | 26 | 23.39 | |
| 0.0884 | 0.0856 | 0.0515 | 0.0915 | 0.2488 | 0.3174 | 0.2898 | 0.3507 | 28 | 22.93 | |

**Fig. 6** Flow diagram of procedure for DEM improvement
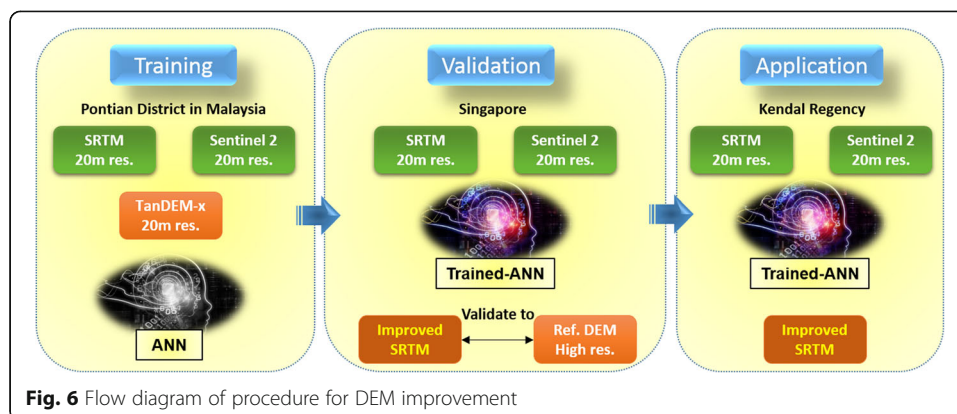
of the region have the same population of L-moments, and then the studied region can be defined as homogeneous or not (Hosking and Wallis 1997). Figure 5 shows the selected 5 grid points which meet the criteria of RFA.

The selected grid points were used to construct the IDF curves for the study area. Design storms, calculated from the IDF curves of different return periods, were then applied to the numerical flood simulation.

## Numerical model setup

The MIKE 21 Flow Model (MIKE 21) is a general numerical modelling system for simulation of water levels and flows (DHI 2017a). This modelling system is applied to coastal, marine engineering, ecological and inland flooding modelling projects. The water levels and flows are calculated on a rectangular grid which covers the area of interest with the information of bathymetry, bed resistance coefficients and hydrographic boundary conditions (Warren and Bach 1992).

The model is divided into two different setups:

- A runoff model domain built with MIKE 11
- 2D hydrodynamic model domain at 20 m resolution due to hefty computational time.

The area for runoff model is mainly mountain area and 2D model is for plain and urban areas. The calculated runoff discharges from MIKE 11 rainfall- runoff model are provided as inputs to the 2D model as its upstream boundary conditions. The division of sub-catchments map and hydrograph of runoff discharge are shown in Fig. 7. The area for runoff model domain is 805.4 km$^2$ and 2D domain is 422.6 km$^2$. Simulation period was calculated from the time of concentration using the velocity method (Heidler 2015) for the study area and the computational time is 10 h per one simulation with 1 s time step. The tide data were extracted from the Global Tide Model (DHI 2017b). The improved SRTM, at 20 m resolution, was used for bathymetry. Table 3 summarizes the details of the model setup.

## Results

### Improved SRTM

The performance of improved SRTM is significant at the validation step in Singapore. The trained ANN using TanDEM-X and Sentinel 2 multispectral imagery was applied
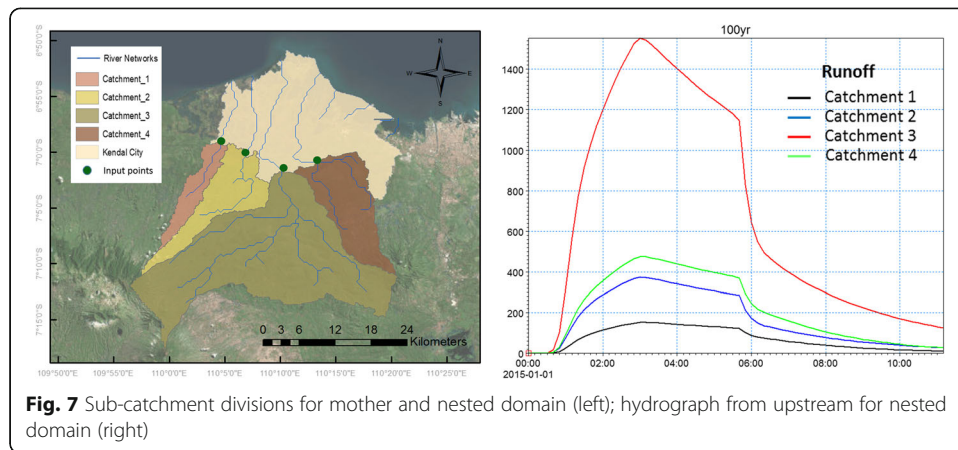
**Fig. 7** Sub-catchment divisions for mother and nested domain (left); hydrograph from upstream for nested domain (right)
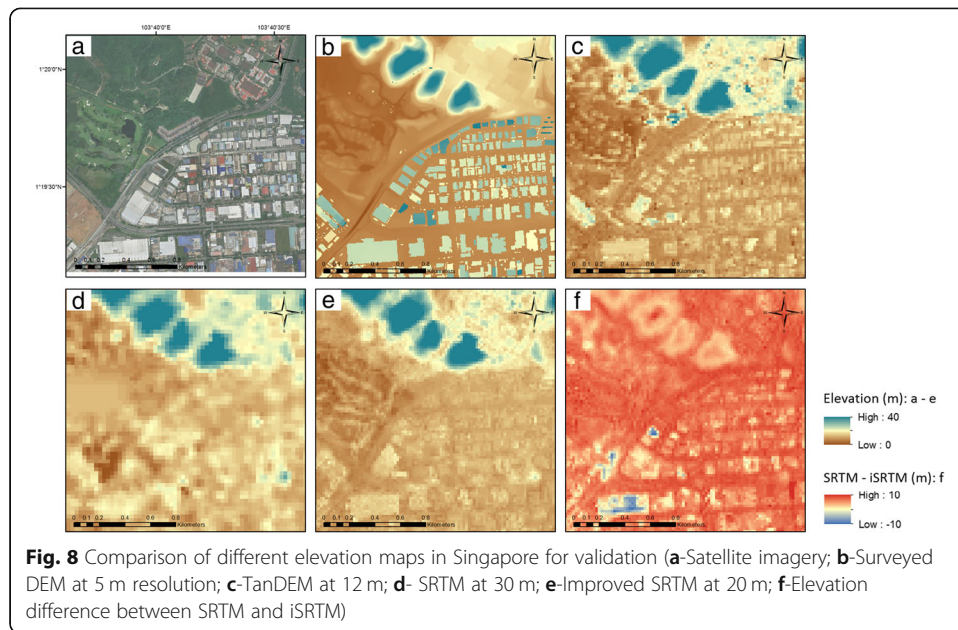
to Singapore area to validate the performance. The original SRTM and improved SRTM were compared to surveyed DEM in Singapore. Figure 8, shows the elevation maps from different sources with a satellite imagery for reference (Fig. 8a). Figure 8b is surveyed DEM at 5 m resolution, (Fig. 8c) is TanDEM-X at 12 m resolution which is used for the target layer, (Fig. 8d) is SRTM at 30 m resolution and (Fig. 8e) is the improved SRTM at 20 m resolution. The improved SRTM shows clearer land shapes and better matches against the elevation of surveyed DEM. Figure 8f shows the difference between SRTM and Improved SRTM. Road areas have relatively larger differences (red color) and the areas with buildings have smaller differences. This implies that the elevation of roads is lowered and buildings are raised so that land shape can be more distinguished in the improved SRTM. Also, there was a significant reduction in the Root Mean Square Errors (RMSE) of 42.3% (7.8 m to 4.5 m) as shown in Fig. 9. This trained ANN was then applied to the study area in Kendal to proceed with flood modelling. As there is no high accuracy DEM to be compared with improved SRTM, the shape of the surfaces was compared against satellite images and original SRTM. The improved SRTM shows clearer topography shapes than original SRTM, especially those river networks, roads and buildings (Fig. 10).

### Design storm scenarios

Meteorological station in Jakarta (Darmaga station), Indonesia with existing IDF curves was compared with the derived IDF curves from WRF/ERA40. The objective is to assess how
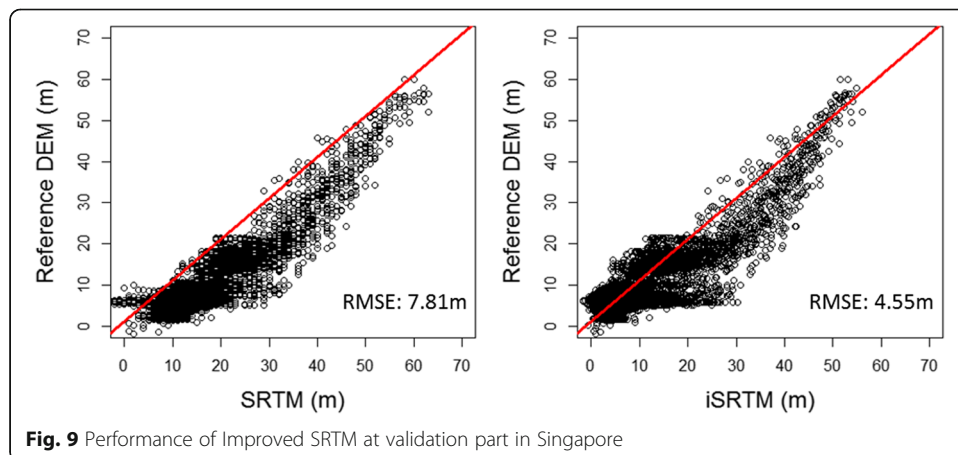
**Table 3** MIKE 21 model setup

| Input type | Remarks |
| --- | --- |
| Catchment areas | 1228 km$^2$ (805.4 km$^2$ for runoff model; 422.6 km$^2$ for 2D model) |
| Simulation time | 11 h |
| Time step | 1 s |
| Tide | Hourly tidal level extracted from Global Tide model |
| Rainfall | IDF Design Storms |
| DEM | 20 × 20 m resolution for 2D model domain<br>Improved SRTM using TanDEM-X + ANN + Sentinel 2 |
| Computational time | 10 h per one simulation |

**Fig. 8** Comparison of different elevation maps in Singapore for validation (**a**-Satellite imagery; **b**-Surveyed DEM at 5 m resolution; **c**-TanDEM at 12 m; **d**- SRTM at 30 m; **e**-Improved SRTM at 20 m; **f**-Elevation difference between SRTM and iSRTM)
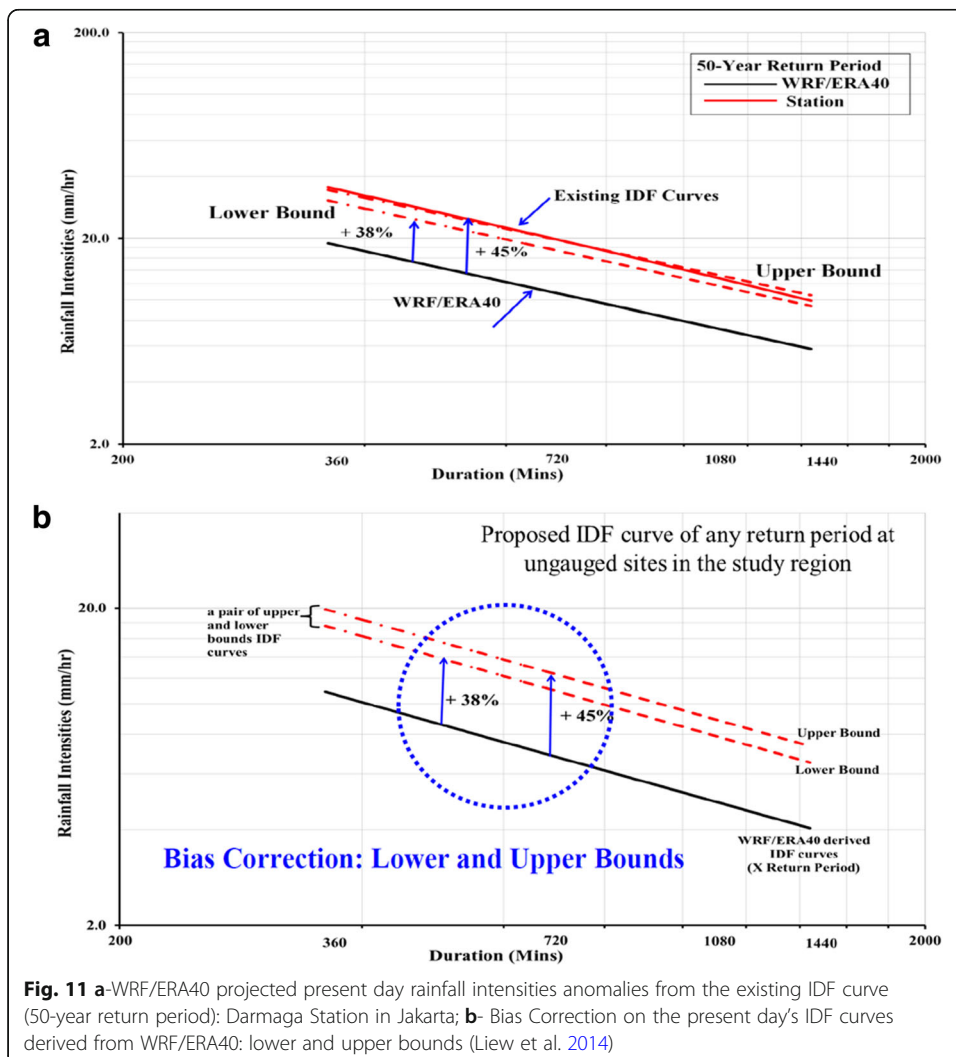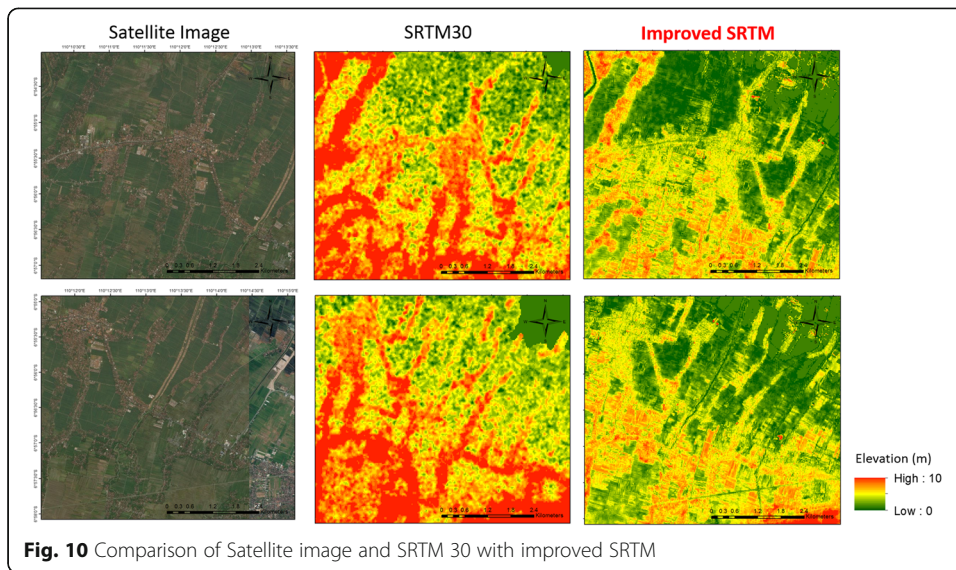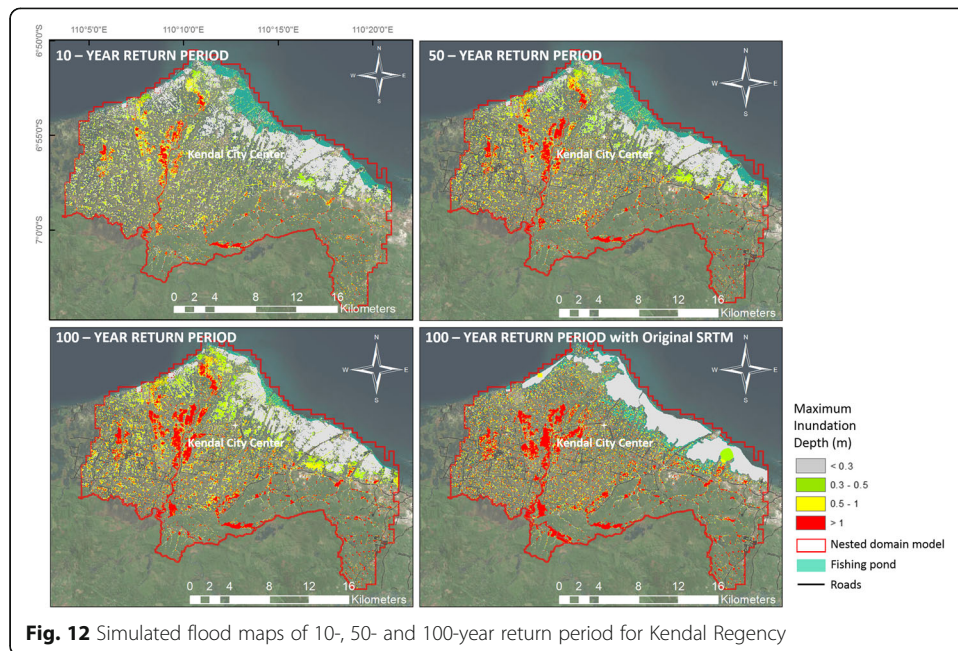
much the WRF/ERA40 derived IDF curves deviate from their observed counterpart. The baseline period considered is from 1961 to 1990. The derived IDF curves consistently underestimate the existing IDF curves within a range of 38% (lower bound) to 45% (upper bound) as shown in Fig. 11a (Liew et al. 2014). This implies WRF/ERA40 has the tendency to underestimate rainfall intensities which serves as a guideline for derivation of present and future IDF curves for sites with short or no rainfall records. This study also adapts the range of bias correction (38% to 45%) to the derived IDF curves as shown in Fig. 11b.

## Flood model simulation results

Flood maps of two scenarios were generated from the M21 numerical model. Figure 12 shows the flood maps of 10-, 50- and 100-year return periods and 100-year return period with original SRTM. The flood map from original SRTM underestimates the flood extent when compared to the improved SRTM. This implies that the improved SRTM has more



**Fig. 9** Performance of Improved SRTM at validation part in Singapore

**Fig. 10** Comparison of Satellite image and SRTM 30 with improved SRTM



**Fig. 11 a**-WRF/ERA40 projected present day rainfall intensities anomalies from the existing IDF curve (50-year return period): Darmaga Station in Jakarta; **b**- Bias Correction on the present day's IDF curves derived from WRF/ERA40: lower and upper bounds (Liew et al. 2014)

**Fig. 12** Simulated flood maps of 10-, 50- and 100-year return period for Kendal Regency

connectivity in the topography allowing the water to flow naturally. Due to the lack of flood information, validating the model cannot be fulfilled in the study. The statistics of flood map extent based on different water depths is shown in Table 4. It can be seen that total inundated areas keep increasing as return period progresses.

The areas of inundation with 0.3–0.5 m water depth are decreasing with higher return period, while the areas with greater than 0.5 m are increasing. 8.3% of the catchment is inundated with 10-year return period and 20% of the area is inundated with 100-year return period. The flooded area of 100-year return period with original SRTM shows less extent than improved SRTM at below 1 m depth due to less connectivity of DEM but it is much higher at above 1 m depth.

## Discussion

The flood hazard assessment has been conducted to address the flooding issues for Kendal Regency where challenges in obtaining high accuracy DEM and good quality with long recorded rainfall data. The improved SRTM was used as the topographic input which generated using ANN with Sentinel 2 multispectral imagery. The ANN was trained in the southern part of Malay Peninsula with TanDEM-X and Sentinel 2. The trained ANN was validated in Singapore and the improved SRTM is marked with a 42.3% RMSE reduction in general. As mentioned in 2.4, ANN was able to classify the
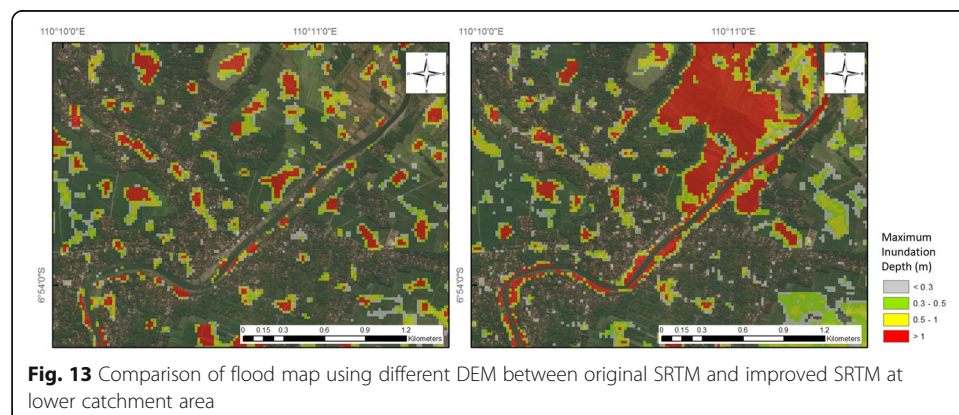
**Table 4** Inundation statistics corresponding to different scenarios for Kendal Regency

| Return periods | 0.3 m – 0.5 m | | 0.5 m – 1 m | | > 1 m | |
|---|---|---|---|---|---|---|
| | km$^2$ | % | km$^2$ | % | km$^2$ | % |
| 10-yr | 13.4 | 3.2 | 12.5 | 3.0 | 8.7 | 2.1 |
| 50-yr | 21.1 | 5.0 | 23.9 | 5.7 | 21.4 | 5.1 |
| 100-yr | 23.6 | 5.6 | 30.2 | 7.1 | 31.0 | 7.3 |
| 100-yr (SRTM) | 14.1 | 3.3 | 22.6 | 5.4 | 38.1 | 9.0 |

land use and cover using the 8 bands of Sentinel 2. Based on the various land characteristics, different weights were calculated to reduce the error between the elevation of SRTM and TanDEM. For example, as shown in (Fig. 8f), the difference between SRTM and Improved SRTM, road areas have relatively larger differences and the areas with buildings have smaller differences. The resolution of DEM was thus able to increase from 30 m to 20 m. This makes the elevation map of the improved SRTM look clearer land shapes than the original SRTM. After the validation, it was applied to Kendal Regency, where surveyed DEM was not available. The improved SRTM and original SRTM for the study area were compared against satellite imagery as reference DEM is not available this area. The shapes of catchment's surface were clearer in improved SRTM, especially river networks, roads and buildings compared to the original SRTM.

The RCM WRF driven by ERA40 produced the mean daily rainfall data from 1961 to 1990 over Southeast Asia domain and have been compared reasonably well with gridded observation data of CRU. The IDF curves then were developed based on the criteria of RFA. It was found that the derived IDF curves underestimate the existing IDF curves (station in Jakarta) within a range of 38% to 45% thus this study proposed to adjust the range of bias correction to the derived IDF curves.

Flood map scenarios were generated using improved SRTM and design storms of 10-, 50- and 100-year return periods using MIKE 21 hydrodynamic model. Two different models were simulated; Rainfall-Runoff model and 2D surface model. Discharges from runoff model were then used as input for 2D model. The generated flood maps and statistic show that the inundated extent area with greater than 0.3 m ranges from 8.3% to 20% of the study area. The flood map of 100-year return periods using original SRTM was compared with using improved SRTM. The flood map with original SRTM has greater area of higher inundation depths (> 1 m) than improved SRTM in (Table 4) and those are concentrated in the middle of catchment (Fig. 12). This implies that due to the limitation of SRTM (i.e. coarse resolution, inaccurate level and less connectivity) the flooded areas are not well distributed compared to the improved DEM. Figure 13 shows enlargement of the 100-yr flood maps in downstream. As the flooded areas are stuck in the middle of catchment from original SRTM, upstream water cannot flow downstream and this resulted in less flooded area in the lower catchment. In addition the flooded areas along the river in improved SRTM look reasonable as these are floodplains with a low-lying ground.



**Fig. 13** Comparison of flood map using different DEM between original SRTM and improved SRTM at lower catchment area

## Conclusions

The approach developed within this project presented an innovative and cost-effective flood hazard assessment. The freely accessible SRTM DEM was improved using remote sensing data and ANN. The developed information is significant and relevant to increase data accuracy allowing use of hydraulic numerical models. RFA was conducted using regional climate model downscaled precipitation data, as proxies, over the study area. The IDF curves with different durations and return periods were constructed as inputs for the numerical model. This method can be applied to areas lacking adequate rainfall records. The methodology and obtained results are significant for smart water management in the areas where data are not sufficient. The paper should be of interest to readers in the areas of remote sensing, artificial intelligence and flood management, especially for the policy makers in proposing relevant flood mitigation measures to the anticipated increasing flood damages, with higher confidence.

### Availability of data and materials
SRTM and Sentinel 2 are publicly accessible through the online. Authors would not share the surveyed DEM and TanDEM-X as those are under the non-disclosure agreement.

### Authors' contributions
Mr. DK; Professor PG and Professor SYL coceived of the main idea for development of improved DEM methodology. DK and SYL desinged the framework and anlaysed the data. DK wrote the manuscript in consulation with PG.  All authors read and approved the final manuscript.

### Competing interests
The authors declare that they have no competing interests.

## Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### References
Abily M, Bertrand N, Delestre O, Gourbesville P, Duluc CM (2016) Spatial global sensitivity analysis of high resolution classified topographic data use in 2D urban flood modelling. Environ Model Softw 77:183–195

Abily M, Delestre O, Amossé L, Bertrand N, Richet Y, Duluc CM, Gourbesville P, Navaro P (2015) Uncertainty related to high resolution topographic data use for flood event modeling over urban areas: toward a sensitivity analysis approach. ESAIM 48:385–399

Bathrellos GD, Karymbalis E, Skilodimou HD, Gaki-Papanastassiou K, Baltas EA (2017) Suitability estimation for urban development using multi-hazard assessment map. Sci Total Environ 575:119–134, ISSN 0048-9697. https://doi.org/10.1016/j.scitotenv.2016.10.025

Bathrellos GD, Skilodimou HD, Chousianitis K, Youssef AM, Pradhan B (2016) Urban flood hazard assessment in the basin of Athens Metropolitan city, Greece. Environ Earth Sci 75(4):319

Coude, R (2018) Background on DEM. Retrieve from http://radiomobile.pe1mew.nl/?Geodata:Background_on_DEM

DHI (2017a) MIKE 21 flow model and MIKE 21 flood screening tool – hydrodyamic module scientific documentation, published by DHI

DHI (2017b) MIKE 21 toolbox – global tide model tidal prediction, published by DHI

Drusch M, Del Bello U, Carlier S, Colin O, Fernandez V, Gascon F, Hoersch B, Isola C, Laberinti P, Martimort P, Meygret A, Spoto F, Sy O, Marchese F, Bargellini P (2012) Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens Environ 120:25–36. https://doi.org/10.1016/j.rse.2011.11.026 ISSN 0034-4257. http://www.sciencedirect.com/science/article/pii/S0034425712000636

Heidler LM (2015) Evaluation of different hydrologycal models in data scarce regions on the island of Ceram, Indonesia, Technische Universitat Munchen, Germany

Hewiston, Crane BC (1994) Precipitation Controls in Southern Mexico, in Neural Nets, Kluwer Academic Publisher

Hosking JRM, Wallis JR (1997) Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge; New York, xiii, p 224

Hsu K, Gupta HV, Sorooshian S (1995) Artificial neural network modeling of the rainfall-runoff process. Water Resour Res 31(10):2517–2530. https://doi.org/10.1029/95WR01955

Jarvis A, Rubiano J, Nelson A, Farrow A, Mulligan M (2004) Practical Use of SRTM Data in the Tropics: Comparisons with Digital Elevation Models Generated from Cartographic Data. Working Document 198. International Center for Tropical Agriculture: Cali, Colombia

Kim D, Sun Y, Wendi D, Jiang Z, Liong SY, Gourbesville P (2018) Flood modelling framework for Kuching City, Malaysia: overcoming the lack of data. In: Gourbesville P, Cunge J, Caignaert G (eds) Advances in Hydroinformatics. Springer Water. Springer, Singapore

Kottek M, Grieser J, Beck C, Rudolf B, Rubel F (2006) World map of the Köppen-Geiger climate classification updated. Meteorol Z 15:259–263. https://doi.org/10.1127/0941-2948/2006/0130

Liew SC, Liu J, Vu MT, Raghavan SV, Nguyen NS, Liong SY (2017) Constructing IDF curves using proxy data from downscaled reanalysis: Kendal regency (Indonesia), Willis Research Network

Liew SC, Raghavan SV, Liong SY (2014) Development of intensity-duration-frequency curves at ungauged sites: risk management under changing climate. Res Lett Geosci Lett 1:8

Liew SC, Raghavan SV, Liong SY (2012) A novel approach, using regional climate model, to derive present and future IDF curves for data scarce sites. IAHR-APD 2012

Manibhushan, Patel N, Sahoo G (2011) Artificial neural network: a tool for classification of land use and land covers using satellite images. Int J Comput Sci Inf Technol 4(2):55–59

Ngoc DV, Gourbesville P (2016) Model uncertainty in flood modelling. Case study at Vu Gia Thu Bon Catchment-Vietnam, Procedia Engineering, vol 154, pp 450–458

Nguyen K, Stewart R, Zhang H, Sahin O (2018) An adaptive model for the autonomous monitoring and management of water end use. Smart Water 3. https://doi.org/10.1186/s40713-018-0012-7

Rodriguez E, Morris CS, Belz JE, Chapin EC, Martin JM, Daffer W, Hensley S (2005) An assessment of the SRTM topographic products, technical report JPL D-31639, Jet Propulsion Laboratory, Pasadena, California, p 143

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 reanalysis. Quart JR Meteorol Soc 131:2961–3012

USGS, 2000, US GeoData digital elevation models, US geological survey, Reston, VA. Available online at: http://erg.usgs.gov/isb/pubs/factsheets/fs04000.html. Accessed 2 Feb 2009

Warren IR, Bach HK (1992) MIKE 21: a modelling system for estuaries, coastal waters and seas. Environ Softw 7(4):229–240, ISSN 0266-9838. https://doi.org/10.1016/0266-9838(92)90006-P

Wendi D, Liong SY, Sun Y, Doan CD (2016) An innovative approach to improve SRTM DEM using multispectral imagery and artificial neural network. J Adv Model Earth Syst 8. https://doi.org/10.1002/2015MS000536

Wessel, B (2016) TanDEM-X ground segment – DEM products specification document, EOC, DLR, Oberpfaffenhofen, Germany, Public Document TD-GS-PS-0021, Issue 3.1. https://tandemx-science.dlr.de/