

MAPPING OF HIGHLY HETEROGENEOUS URBAN STRUCTURE TYPE FOR FLOOD VULNERABILITY ASSESSMENT

T. H. Tam¹, M. Z. A. Rahman^{1,*}, S. Harun², I. U. Kaoje¹

¹ TropicalMap Research Group, Universiti Teknologi Malaysia, Johor Bahru, Malaysia – thtam2@live.utm.my

² Dept. of Hydraulic and Hydrology, Universiti Teknologi Malaysia, Johor Bahru – sobriharun@utm.my

KEY WORDS: Urban Structure Type, Object-based image analysis, Random forest, Support vector machine, CART, Flood vulnerability

ABSTRACT:

Vulnerability plays an important role in risk assessment. For flood vulnerability assessment, the map and characteristics of elements-at-risk at different scales are strongly required depending on the risk and vulnerability assessment requirements. This study proposes a methodology to classify urban structure type by combining object-based image classification and different high resolution remote sensing data. In this study, a high resolution satellite image and LiDAR have been acquired over Kota Bharu, Kelantan which consists of highly heterogeneous urban structure type (UST) classes. The first stage is data pre-processing that includes orthorectification and pansharpener of Geoeye satellite image, image resampling for normalised Digital Surface Model (nDSM) and followed by image segmentation for creating meaningful objects. The second stage comprises of derivation of image features, generation of training and testing datasets, and classification of UST. The classification was based on three types of machine learning classifiers, i.e. Random Forest (RF), Support Vector Machine (SVM) and Classification and Regression Tree (CART). The results obtained from the classification processes were compared using individual omission and commission error, overall accuracy and Kappa coefficient. The results show that Random Forest classifier with all image features achieved the highest overall accuracy (93.5%) and Kappa coefficient (0.94). This is followed by CART classifier with overall accuracy of 93.7% and Kappa coefficient of 0.92. Finally, SVM classifier produced the lowest overall accuracy and Kappa coefficient with 88.6% and 0.86, respectively. The UST classification result can be further used to assist detailed building characterisation for large scale flood vulnerability assessment.

* Corresponding author

1. INTRODUCTION

1.1 Introduction

In the context of flood management, flood risk assessment have widely been implemented to replace the traditional flood management. Details about the flood risk assessment can be referred to EU Flood Directive (2007/60/EC). Its aim is to reduce and manage the risks that floods pose to human health, the environment, cultural heritage and economic activity. The Directive requires Member States to first carry out a preliminary assessment by 2011 to identify the river basins and associated coastal areas at risk of flooding. For such zones they would then need to draw up flood risk maps by 2013 and establish flood risk management plans focused on prevention, protection and preparedness by 2015.

According to Penning-Rowsell et al. (2005) and Foudi and Oses-Eraso (2014), the method used to assess flood-risk consists of four steps: (i) hazard assessment; (ii) exposure assessment; (iii) vulnerability assessment; and (iv) risk assessment. Among these four components of flood risk assessment, remote sensing imagery have extensively contributed in hazard assessment to achieve different purposes and scales. For instance, satellite-based rainfall estimates for estimating maximum discharge. A number of studies have been carried out using various satellite-based rainfall estimates (TRMM, CMORPH and PERSIANN) in which results show that estimates to generate streamflow as reliable as using rain gauge data, provided those estimates have to be calibrated (Stisen, S. and Sandholt, I. 2010; Meng et al. 2014; Zubieta et al. 2015; Jiang et al. 2010). In addition to estimate maximum discharge, producing flooded area is of great importance in flood hazard for which different resolution of Digital Elevation Model (DEM) is created using remote sensing technology. LiDAR data with 1 meter high resolution, higher accuracy of flood map can be produced using 1D hydrodynamic model or 2D hydrodynamic model, Such studies have been carried out by Costabile et al. (2015), McDougall and Temple-Watts (2012), and Turner et al. (2013).

For vulnerability assessment, remote sensing data often to provide some biophysical parameters for mapping of flood vulnerability map in qualitative way using multi-criteria methods (e.g. Analytic hierarchy Process, AHP) (Rimba et al. 2017; Ouma and Tateishi, 2014; Sambah and Miura, 2014) or vulnerability index (Kumar and Kunte, 2012; Islam et al. 2016). In this paper, we adopted the definition of vulnerability by Merz et al. (2007) where vulnerability is composed of two elements, exposure and susceptibility. However, less study focus on exposure assessment using remote sensing data (Gerl et al., 2014; Angela et al. 2013). According to Merz et al. (2007), exposure can be quantified by the number of the value of elements which are at risk. The term *elements at risk* includes all elements of the human system, the built environment and the natural environment that area at risk of flooding in a given area. In this study, we focus elements at risk on built environment which can be described by land use land cover.

Land use land cover map play an important role in flood risk assessment. Land use land cover maps describe the spatial location of variety type of elements at risk in the flooded area. However, existing coarser land use classes like residential buildings, industry, and agriculture are often obtained from agency and integrated in flood loss modelling (Foudi et al 2015; Gain et al. 2015; Budiyo et al. 2015; Cammerer et al. 2013). Some studies focused on micro-scale, however, most of the

adopted land use data is relatively generalized (Velasco et al. 2015). In Malaysia, Department of Irrigation and Drainage (DID) produced a guidelines for assessment of flood damage in which the adopted land use information in this guideline is rather generalized (report: *Guidelines and Procedures for the Assessment of Flood Damages in Malaysia*, 2003). Therefore, it is timely to update the detailed land use land cover information to assess the current flood risk situation in Malaysia and subsequently it facilitate related-agency to reduce flood damage. With the advent of remote sensing technology with the aid of image processing and classification method, such objective can be easily achieved, referred to (Manandhar et al. 2009; Adam et al. 2013; Yousefi et al. 2015; Mahmon et al. 2015).

A more detailed land use land cover is refer to classification of Urban Structure Type (UST). Urban structure types is defined as spatial units of the built environment at an aggregated neighbourhood scale, which are relatively homogeneous with respect to the overall population in terms of their physical appearance (land cover) and usage (land use) (Wieland et al., 2016). According to Wieland et al. (2016), characterization and analysis of an exposed building on the basis of its susceptibility attributes and combine with the output of hazard analysis is used to compute the expected damage. Several studies exist that use high resolution image with object-based image analysis for the classification of UST where these studies mostly at Germany in which the urban structure type is well planning (voltersen et al. 2014; Puissant et al. 2012; Bochow et al. 2010; Walde et al. 2014). However, Bochow et al. (2010) tried to classify slum areas as one of the urban structure types using Maximum Likelihood as classifier. Result shows highest error in term of commission error and omission error.

Based on previous studies there are still lack of effort to classify highly heterogeneous UST in developing countries using high resolution remote sensing data. This has definitely complicate the process of flood vulnerability and risk assessment, especially in developing countries where detailed elements-at-risk characteristic should be defined prior to vulnerability assessment. The aim of this study is to compare different methods of machine learning classifier in producing urban structure type map using combining high resolution satellite image and airborne LiDAR data.

2. MATERIALS

2.1 Description of Study area

The city of Kota Bharu, Kelantan is located in north eastern part of Peninsular Malaysia where its neighbour country is Thailand. The Kota Bharu is the state capital and has become the central point for Kelantan's administrative and business activities. The landscape of Kota Bharu is different to other cities in Malaysia in which surrounded by a number of Malay villages also known as *kampung*. According to Fee (1998), the development of kampung does not comply any plan, therefore, the spatial fabric of kampung is highly heterogeneity. The studied area shown in Figure 1 is part of city of Kota Bharu. Traditional houses (Kampung house) at Kota Bharu are mostly elevated wooden houses. Nowadays, most of these traditional houses are converted into different type of houses (Table 1).

The topography of Kota Bharu city is relatively flat with average elevation of 5 m, please refer to Figure 2a. Every year, the Northeast Monsoon from November to January brought heavy rainfall to Kota Bharu which will cause flooding. Due to this reason, the traditional houses at this city are elevated house.

2.2 Description of Data

There are two main data used in this study, i.e. GeoEye-1 (GE1) satellite image, and airborne Lidar data. Geoeeye-1 satellite image was captured on June 6, 2016. The GE1 data come with four, 2.0 m spatial resolution, multispectral bands (Blue, Green, Red, and NIR) and one, 0.5 m resolution, panchromatic band. The complete LiDAR data (DSM and DEM) with 2 m spatial resolution was obtained in 2008. Both dataset have 8 years difference, however, due to slow development in Kota Bharu, the LiDAR data is still applicable to be used in this study.

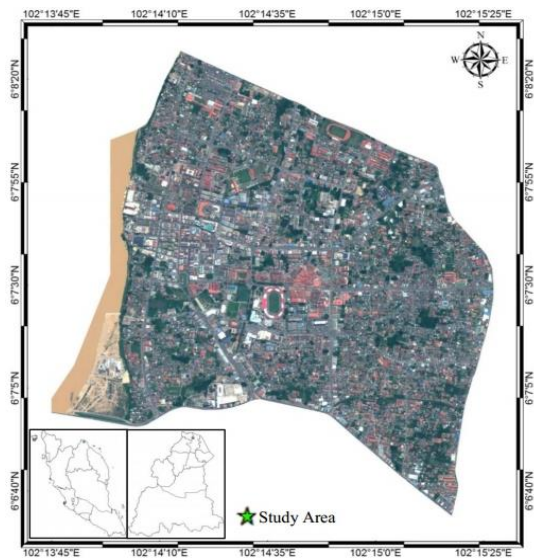


Figure 1. Study area

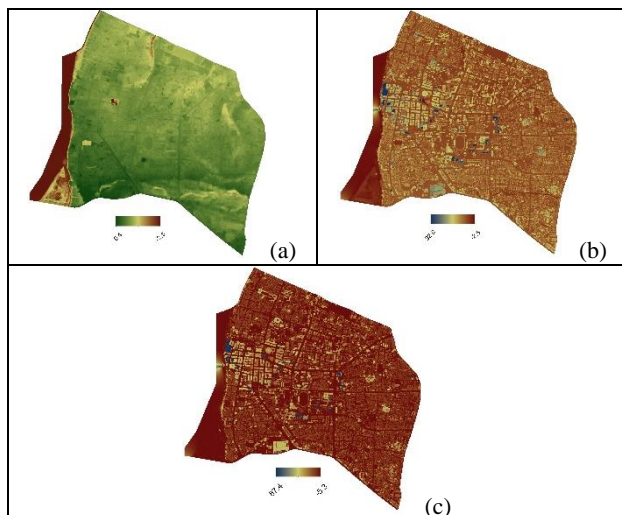


Figure 2. airborne LiDAR data (a) DTM, (b) DSM, (c) nDSM

2.3 Method

In this section, the flowchart of this study is divided into two parts, i.e. data pre-processing and classification of UST with assess its reliability (Figure 3).

Table 1. Different type of houses.

- 1 Single storey concrete house with tile roofing
- 2 Single storey concrete house with metal roofing
- 3 Modified single storey concret house (partial second floor additions - concrete wall) with tite roofing
- 4 Modified single storey concret house (partial second floor additions - concrete wall) with metal roofing
- 5 Modified single storey concret house (partial second floor additions - wooden wall) with tite roofing
- 6 Elevated wooden house with metal roof (zinc)
- 7 Elevated wooden house with tile roofing
- 8 Elevated wooden and partial concrete house with metal roofing (zinc)
- 9 Elevated wooden house and partial concrete with metal roof (zinc)
- 10 Elevated concrete house with tile roofing
- 11 Elevated concrete house with metal roofing
- 12 Wooden house with metal roofing (zinc)
- 13 Double storey concret house with tile roofing
- 14 Double storey concret house with metal roofing
- 15 Double storey (ground floor: concrete; first floor: wood) with metal roofing
- 16 Double storey (ground floor: concrete; first floor: metal) with metal roofing
- 17 Mix single and double storey concrete house with metal/tile roofing
- 18 Mix elevated wooden and partial single storey concrete house (metal and tile roofing)
- 19 Mix elevated and partial single storey concrete house with tile roofing
- 20 Mix elevated wooden and partial single storey concrete house with metal roofing
- 21 Mix wooden and concrete house with metal roofing

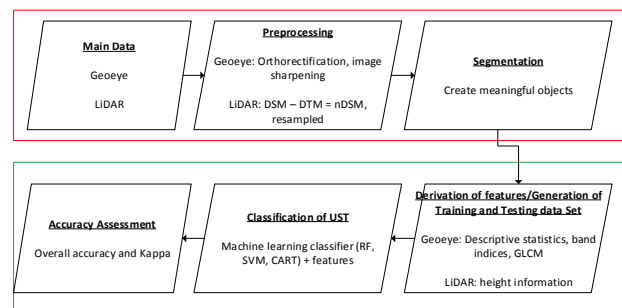


Figure 3. Flow chart, red box is data pre-processing, while green box is classification of UST.

2.3.1 Data pre-processing: Orthorectification and image sharpening applied to GE-1 satellite image. An nDSM is created by subtracting Digital Terrain Model (DTM) from a Digital Surface Model (DSM). Next, resampling the nDSM to 0.5 m in order to match the resolution of satellite image. Prior to run image segmentation, a tool called ‘Estimation of Scale Parameters’ (ESP) version 2 developed by Drägut et al. (2010) was used to estimate the suitable scale parameter for multiresolution image segmentation. The scale parameter estimated by ESP is 116 and associated with other segmentation parameters, i.e. shape (0.9) and compactness (0.1) are well used for creating meaningful objects in this study.

2.3.2 UST mapping: In this step, the classification of UST is important. As there are no compulsory and consistent urban structure types, the classification scheme depends on the study area and the objective of the respective analysis (Banzhaf & Höfer, 2008). Therefore, we defined 8 USTs which shown in Figure 4.

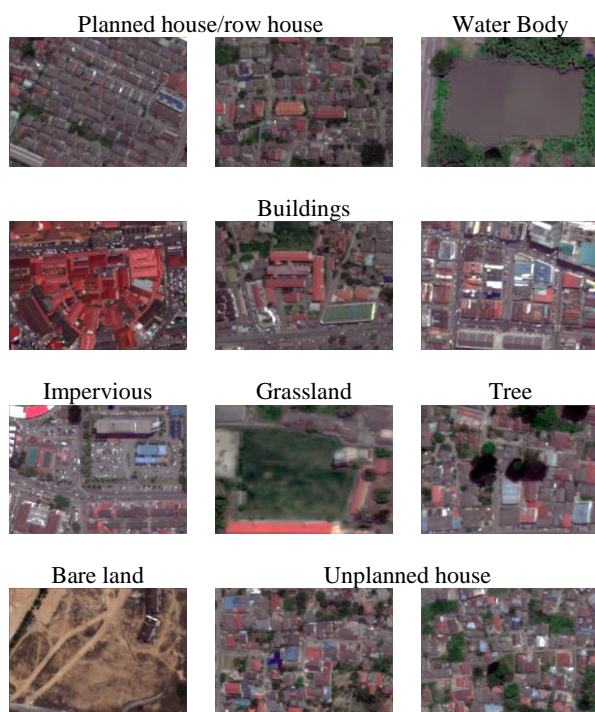


Figure 4. The urban structure type at study area

In this study, we derived a total of 74 image features on the basis of four feature classes, i.e. Spectral, Band indices, Haralick texture, and height information (Table 2). In addition, three different machine learning classifiers, i.e. Random Forest (RF), Support Vector Machines (SVM) and Classification and Regression Tree (CART), were used to implement UST classification. However, present study is an preliminary study on the classification of UST, therefore, the parameter of each classifier is set to default.

Both CART and Random Forest are tree-based algorithm. CART is a single tree of classification model in which we can prune the overgrown trees while RF is an ensemble of number of CART. Also, pruning is not allowed in it thought some tuning function is available in which we can pass parameter like no of trees etc. SVM is a non-parametric algorithm that allow fast and dependable classification with limited amount of data.

Following an object-based approach to image analysis, segments have been used as sample type. In this study, total of 460 segments were generated. Total of 36% of the segments are assigned as training data sets while the rest are testing data sets (295 segments). Training sets data are used for classification purposes while testing sets data are used to evaluate the accuracy of classification result.

Next, accuracy assessment of image classificaiton is carried out by calculation of overall accuracy and Kappa coefficient which are the standard and common method to assess the reliability of the produced thematic map. According to Thomlinson et al. (1999) accurate a classification when the OA is at least equal to 85% and no class is less than 70%.

3. RESULTS AND DISCUSSION

This section briefly explained on the results of UST classification shown in Figure 5 and discuss the performance of every machine learning classifier, the producer’s accuracy and user’s accuracy of every class and also the influence of the image features used for classification.

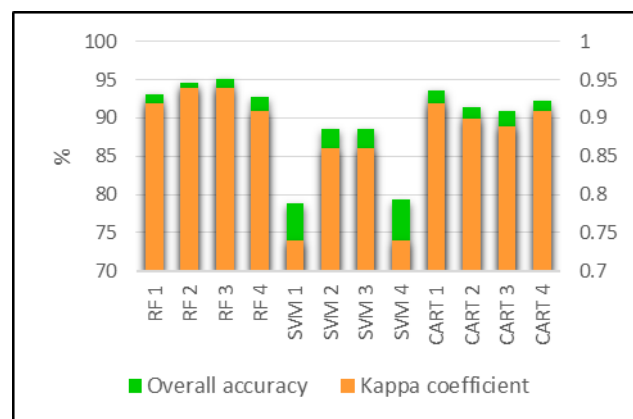


Figure 5. Comparison of every classifier in term of overall accuracy and Kappa coefficient.

Classification of UST was conducted using three different machine learning classifiers, which were CART, RF and SVM. The classification accuracy statistics are summarized in Figure 4. The results show that classification using RF3 provided the highest overall accuracy (95.2%) and Kappa coefficient (0.94). SVM generated the least accurate classification map with 78.9% overall accuracy and Kappa coefficient of 0.74. Moreover, CART produced relatively high overall accuracy and Kappa coefficient with above than 90 % and 0.85, respectively. The reason SVM produced least overall accuracy compared to other two classifiers is attributed to the parameter of SVM is set to default.

For the RF3, the classes with the highest producer’s accuracy were those of water (100%), tree (100%), planned house (100%), impervious (100%), and grassland (100%) followed by unplanned house (97.8%), bare land (94.7%) and buildings (83.6%). User’s accuracy was higher for water (100%), Tree (100%), planned house (100%), impervious (100%), and grassland (100%), followed by buildings (98.6%), bare land (96.6%), and unplanned house (79.1%). All classes were easily separable by all classifier algorithms applied.

Table 3. Summary of confusion matrix of each finest classifier.

Class	CART 1		SVM 2		RF 3	
	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.
Water	100.0	99.0	100.0	100.0	100.0	100.0
Unplanned house	88.1	93.7	81.3	78.0	97.8	79.1
Tree	98.9	96.9	93.3	90.1	100.0	100.0
Planned house	100.0	100.0	0.0	0.0	100.0	100.0
Impervious	100.0	100.0	100.0	100.0	100.0	100.0
Grassland	95.9	97.3	86.5	90.7	100.0	100.0
Buildings	88.7	92.1	86.4	84.3	83.6	98.6
Bare land	88.8	66.4	87.6	100.0	94.7	96.6

In addition, we divided the 74 sub features into four groups (Table 2) in order to examine their influence on classification. Apparent result of classification using RF classifier depicted that taking more sub-features into consideration, the overall accuracy and Kappa coefficient is increasing. Similar condition is noticeable in using SVM classifier. On the contrary, the overall accuracy and Kappa coefficient of CART classifier is decreasing with increase the numbers of sub-features.

Table 2. The main feature class and its associated image features.

Image Feature description	Feature class	G1	G2	G3	G4
Mean DN value in image band i	Spectral	x	x	x	x
Standard DN of reflectance values in image band i	Spectral	x	x	x	x
Maximum of DN value in image band i	Spectral	x	x	x	x
Minimum of DN value in image band i	Spectral	x	x	x	x
Median of DN value in image band i	Spectral	x	x	x	x
Mean, standard deviation, median and mode values of normalized difference vegetation index (NDVI)	Band index	x	x	x	x
Mean, standard deviation, median and mode values of soil adjusted vegetation index (SAVI)	Band index	x	x	x	x
Mean, standard deviation, median and mode values of simple ratio (SR)	Band index	x	x	x	x
Mean and standard deviation of height values	Elevation			x	x
Angular second moment derived from the grey-level co-occurrence matrix (GLCM) in band x	Textural		x	x	
Homogeneity derived from GLCM in band x	Textural		x	x	
Contrast derived from GLCM in band x	Textural		x	x	
Dissimilarity derived from GLCM in band x	Textural		x	x	
Entropy derived from GLCM in band x	Textural		x	x	
Angular 2nd moment derived from GLCM in band x	Textural		x	x	
Mean derived from GLCM in band x	Textural		x	x	
Standard deviation derived from GLCM in band x	Textural		x	x	
Correlation derived from GLCM in band x	Textural		x	x	

The difference of group 1 and group 4 of sub-features is group 4 take height information into consideration. By looking at classification accuracy statistics, it noticed that the overall accuracy of group 4 is lower than group 1. The overall accuracy of RF1 and RF4 are 93.2% and 92.8%, respectively where the overall accuracy of CART1 (93.7) and CART 4 (92.3%). The reason is that the Geoeye-1 satellite image captured at 19.3° off nadir. However such off nadir degree is considerable large compared to LiDAR data, therefore, this is the source of causing misclassification.

4. CONCLUSION

In this study, classification of UST by using three different machine learning classifiers is conducted. Results show that Random Forest classifier achieve highest overall accuracy of 95.2% and Kappa coefficient of 0.94, followed by CART classifier and SVM classifier. Nevertheless, further study should be carry out in few aspects. The building class in present study is too general, further study should try to classify different type of buildings. In addition, more focusing on tuning parameter of classifier, especially SVM. This result can be further use to assist detailed building characterisation for flood vulnerability assessment.

ACKNOWLEDGEMENTS

The authors are grateful to the Universiti Teknologi Malaysia (UTM) for the financial support under grant GUP Q.J130000.2527.17H76.

REFERENCES

- Adam, E., Mutanga, O., Odindi, J. and Abdel-Rahman, E.M., 2013. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, 35(10), pp. 3440-3458.
- Angela, B.V., Norbert, H. and Jochen, S., 2013, April. Building extraction from remote sensing data for parameterising a building typology: a contribution to flood vulnerability assessment. In *Urban Remote Sensing Event (JURSE)*, 2013 Joint (pp. 147-150). IEEE.
- Bochow, M., Taubenböck, H., Segl, K. and Kaufmann, H., 2010, July. An automated and adaptable approach for characterizing and partitioning cities into urban structure types. In *Geoscience and Remote Sensing Symposium (IGARSS)*, 2010 IEEE International (pp. 1796-1799). IEEE.
- Budiyono, Y., Aerts, J., Brinkman, J., Marfai, M.A. and Ward, P., 2015. Flood risk assessment for delta mega-cities: a case study of Jakarta. *Natural hazards*, 75(1), pp. 389-413.
- Cammerer, H., Thieken, A.H. and Verburg, P.H., 2013. Spatio-temporal dynamics in the flood exposure due to land use changes in the Alpine Lech Valley in Tyrol (Austria). *Natural Hazards*, 68(3), pp. 1243-1270.
- Costabile, P. and Macchione, F., 2015. Enhancing river model set-up for 2-D dynamic flood modelling. *Environmental Modelling & Software*, 67, pp. 89-107.
- Fee, C. V. 1998. *The Encyclopedia of Malaysia: Architecture*. Archipelago Press, Kuala Lumpur, pp. 125.
- Foudi, S., Osés-Eraso, N. and Tamayo, I., 2015. Integrated spatial flood risk assessment: The case of Zaragoza. *Land Use Policy*, 42, pp. 278-292.
- Gain, A.K., Mojtahed, V., Biscaro, C., Balbi, S. and Giupponi, C., 2015. An integrated approach of flood risk assessment in the eastern part of Dhaka City. *Natural Hazards*, 79(3), pp. 1499-1530.
- Gerl, T., Bochow, M. and Kreibich, H., 2014. Flood damage modeling on the basis of urban structure mapping using high-resolution remote sensing data. *Water*, 6(8), pp.2367-2393.
- Islam, M.A., Mitra, D., Dewan, A. and Akhter, S.H., 2016. Coastal multi-hazard vulnerability assessment along the Ganges deltaic coast of Bangladesh—A geospatial approach. *Ocean & Coastal Management*, 127, pp. 1-15.
- Jiang, S.H., Ren, L.L., Yong, B., Yang, X.L. and Shi, L., 2010. Evaluation of high-resolution satellite precipitation products with surface rain gauge observations from Laohahe Basin in northern China. *Water Science and Engineering*, 3(4), pp. 405-417.
- Kumar, A.A. and Kunte, P.D., 2012. Coastal vulnerability assessment for Chennai, east coast of India using geospatial techniques. *Natural Hazards*, 64(1), pp. 853-872.
- Mahmon, N.A., Ya'Acob, N. and Yusof, A.L., 2015, March. Differences of image classification techniques for land use and land cover classification. In *Signal Processing & Its Applications (CSPA)*, 2015 IEEE 11th International Colloquium on (pp. 90-94). IEEE.
- Manandhar, R., Odeh, I.O. and Ancev, T., 2009. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sensing*, 1(3), pp. 330-344.
- McDougall, K. and Temple-Watts, P., 2012. The use of LiDAR and volunteered geographic information to map flood extents and inundation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, pp. 251-256.
- Meng, J., Li, L., Hao, Z., Wang, J. and Shao, Q., 2014. Suitability of TRMM satellite rainfall in driving a distributed hydrological model in the source region of Yellow River. *Journal of Hydrology*, 509, pp. 320-332.
- Merz, B., Thieken, A.H. and Gocht, M., 2007. Flood risk mapping at the local scale: concepts and challenges. In *Flood risk management in Europe* (pp. 231-251). Springer, Dordrecht.
- Ouma, Y.O. and Tateishi, R., 2014. Urban flood vulnerability and risk mapping using integrated multi-parametric AHP and GIS: methodological overview and case study assessment. *Water*, 6(6), pp. 1515-1545.
- Penning-Rowsell, E., Johnson, C., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J. and Green, C., 2005. The benefits of flood and coastal risk management: a handbook of assessment techniques. ISBN 1904750516.
- Puissant, A., Zhang, W. and Skupinski, G., 2012, May. Urban morphology analysis by high and very high spatial resolution remote sensing. In *International Conference on Geographic Object-Based Image Analysis* (Vol. 4, pp. 524-529).
- Rimba, A.B., Setiawati, M.D., Sambah, A.B. and Miura, F., 2017. Physical flood vulnerability mapping applying geospatial techniques in Okazaki City, Aichi Prefecture, Japan. *Urban Science*, 1(1), p.7.
- Rimba, A.B., Setiawati, M.D., Sambah, A.B. and Miura, F., 2017. Physical flood vulnerability mapping applying geospatial techniques in Okazaki City, Aichi Prefecture, Japan. *Urban Science*, 1(1), p. 7.
- Stisen, S. and Sandholt, I., 2010. Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling. *Hydrological Processes*, 24(7), pp. 879-891.
- Turner, A.B., Colby, J.D., Csontos, R.M. and Batten, M., 2013. Flood modeling using a synthesis of multi-platform LiDAR data. *Water*, 5(4), pp. 1533-1560.

Velasco, M., Cabello, À. and Russo, B., 2015. Flood damage assessment in urban areas. Application to the Raval district of Barcelona using synthetic depth damage curves. *Urban Water Journal*, 13(4), pp.426-440.

Voltersen, M., Berger, C., Hese, S. and Schmullius, C., 2014. Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level. *Remote Sensing of Environment*, 154, pp.192-201.

Walde, I., Hese, S., Berger, C. and Schmullius, C., 2014. From land cover-graphs to urban structure types. *International Journal of Geographical Information Science*, 28(3), pp. 584-609.

Wieland, M., Torres, Y., Pittore, M. and Benito, B., 2016. Object-based urban structure type pattern recognition from Landsat TM with a Support Vector Machine. *International Journal of Remote Sensing*, 37(17), pp. 4059-4083.

Yousefi, S., Khatami, R., Mountrakis, G., Mirzaee, S., Pourghasemi, H.R. and Tazeh, M., 2015. Accuracy assessment of land cover/land use classifiers in dry and humid areas of Iran. *Environmental monitoring and assessment*, 187(10), p.641.

Zubieta, R., Getirana, A., Espinoza, J.C. and Lavado, W., 2015. Impacts of satellite-based precipitation datasets on rainfall-runoff modeling of the Western Amazon basin of Peru and Ecuador. *Journal of Hydrology*, 528, pp. 599-612.