Mindanao Journal of Science and Technology Vol. 17 (2019) 214-226

Delineation of Flood-Prone Areas in Data-Scarce Environment Using Linear Binary Classifiers

Wendell D. Talampas* and Dennis A. Tarepe
College of Science and Mathematics
University of Science and Technology of Southern Philippines – Cagayan de Oro
Cagayan de Oro City, 9000 Philippines
**wendell.talampas@ustp.edu.ph*

Date received: March 5, 2019
Revision accepted: June 5, 2019

Abstract

The knowledge and understanding of areas prone to flooding are critical in mitigating and reducing risks caused by extreme water events. In the Philippines, however, large scale riverine flooding still remains a huge challenge in many parts of the country as most river basins are still ungauged. This limitation has consequently led to data scarcity and deficiency and has prevented the application of comprehensive floodplain mapping techniques. This study aims to address such limitations by introducing a simplified method for the preliminary identification of flood exposed areas within a river system using geomorphic classifiers. The Geomorphic Flood Index (GFI), a linear binary classification technique, is used in delineating flood-prone areas in a watershed within the Cagayan de Oro River Basin. The index used three morphological features – the drainage network, the contributing area, and the water level to perform the linear binary classification. Results of the study showed the potential of the GFI in delineating flood-prone areas in data-scarce environments particularly in the watershed of the Cagayan de Oro River.

Keywords*:* GFI, linear binary classifiers, flood-prone area, image classification, confusion matrix

---

1. Introduction

One research topic of increasing interest is the determination of floodplain areas in data-scarce environments. There had been several research studies conducted on addressing flood hazard prediction. Some of these studies addressed the prediction of floods at ungauged sites by operating a transfer hydrologic information using regionalization methods (Bloschl and Sivaplan, 1997; Merz and Bloschl, 2008; Padi *et al.*, 2011). Others proposed simplified global-scale models of surface water flows based on hydrological routing schemes driven regional or global climate models (Herold and Mouton, 2011;

Pappenberger *et al.*, 2012). Hydraulic models, however, have drawbacks being computationally intensive, costly and difficult to implement over large study areas. Consequently, hazard assessment over large, unstudied areas poses a significant challenge. Therefore, there is a need to look for alternative ways of generating flood inundation maps in an effective and inexpensive means. Extensive investigations using different hydrologic, climatic and topographic features have already been conducted in several Italian gauged basins (Manfreda *et al.*, 2014); an ungauged basin in Africa (Manfreda *et al.*, 2014) and over the entire continental U.S. (Samela *et al.*, 2017) to identify good indicators of flood hazard exposure. Based on these extensive investigations, a method called Geomorphic Flood Index (GFI), based on linear binary classifiers has consistently exhibited higher classification accuracies compared to other candidates. It presented low sensitivity to changes in the input data in terms of dominant topography of the training area, size of the training area, Digital Elevation Model (DEM) resolution, standard flood maps adopted, return time and the scale of the analysis. This study investigated the delineation of flood-prone areas in one of the watersheds of the Cagayan de Oro River Basin using the GFI method.

2. Methodology

The city's intensive developments and rapid urbanization increase the occurrence of riverine flooding and other extreme water events. Since the December 2011 flood that claimed thousands of lives and damaged several million of properties, the city has been preparing for similar extreme water events. The December 2012 and December 2017 flood incidents brought about the Typhoons Pablo and Vinta, respectively, were among the worst floods that hit the city. This extreme water event can be attributed to Cagayan de Oro River Basin's inherent features that made it naturally liable to flooding which is further exacerbated by anthropogenic activities. Knowledge and understanding of flood-prone areas within the river basin are essential in mitigating and reducing the loss of lives and damage to properties. However, in a data-scarce environment, the need to identify these areas using the least amount of data is of imperative concern. The use of linear binary classification in delineating flood-prone areas in the Cagayan de Oro River Basin is an attempt to address such concern.

The delineation of flood-prone and non-flood prone areas in this study were carried out using linear binary classification. The study employed ArcGIS and MATLAB for the pre-processing, processing and post-processing of datasets.

2.1 Digital Elevation Models

For the purpose of this study, DEMs from the U.S. Geological Survey (USGS) Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS) were used. The 3-arc second resolution void-filled DEM (DEM-void) and the hydrologically conditioned DEM (DEM-con) were particularly used and regridded to 90-meter cell size.

2.2 Standard Flood Hazard Map

A 100-year standard flood map for Cagayan de Oro City derived from the University of the Philippines Disaster Risk and Exposure Assessment for Mitigation (UP DREAM) Project was used for training the classifiers. To be able to use the standard flood map, it was converted into a raster file and reclassified to represent flooded areas with a class value of 1 and not flooded areas with a class value of 2. The resulting reclassified flood map was used to generate reference points as ground truths for the classification process.

2.3 DEM Extraction and Pre-Processing for the Area

Using ArcGIS, void-filled and hydrologically-conditioned DEM for the study area were extracted from the HydroSHEDS DEM. Flow direction and flow accumulation were then generated from these DEMs. Generated raster files including the void-filled and hydrologically-conditioned DEMs were then converted to ASCII files for processing into MATLAB.

2.4 Linear Binary Classification

Using MATLAB, a linear binary classification was performed using the input datasets to identify flooded and not flooded areas. For this study, three morphological features namely the drainage network, the contributing area, and the water level were utilized to perform the linear binary classification. These features were utilized as they relate to some of the morphological characteristics of a basin to estimate the water depth calculated as a function of the contributing area and take into account the essential role assumed by the flow depth on floodplain delineation (Samela *et al.*, 2017). For this study, the morphological descriptor GFI was defined as,

$$In\left(\frac{h_r}{H}\right) \tag{1}$$

The GFI compares each point of the basin variable water depth h_r with the elevation difference H. The variable water depth h_r, is computed as a function of the contributing area in the nearest point of the drainage network hydrologically connected to the point under examination,

$$h_r \approx bA_r^n \tag{2}$$

As the calculation takes into account an estimate of the water level in the nearest element of the drainage network, the nearest river was then considered as a source of hazard. The resulting image (GFI image) from the MATLAB processing was then manipulated in the ArcGIS environment for post-processing and evaluation of results.

2.5 Image Classification

Using the resulting image from MATLAB, flooded and not flooded areas were identified and classified and coded as 1 for flooded and 2 for not flooded. To proceed with the image classification, accuracy assessment using supervised image classification was used with standard flood maps as training sites. The intent was to categorize all pixels in the standard flood map into one of several classes or themes. By classifying the image (standard flood map), the study was able to identify and portray the features occurring in the image in terms of the classes or themes these features actually represented in the ground. Using supervised image classification, the study identified flooded and not flooded areas from the standard flood map.

These training sites or test pixels were represented by reference points distributed evenly throughout the map where flooded and not flooded areas are present. Since there were only two classes involved – flooded and not flooded, the rule of thumb was to have ten times the number of test points for each class. Thus, the test points for each class is twenty. A number code was assigned for each class (e.g. 1 for flooded areas; 2 for not flooded). Using the Geoprocessing tool of ArcGIS, the system was set to align the pixels to be created from the reference points with the pixels of the classification as shown in Figure 1. Under the environment setting, the extent and snap raster were set as that of the classified image. The pixels should align correctly with the reference points.

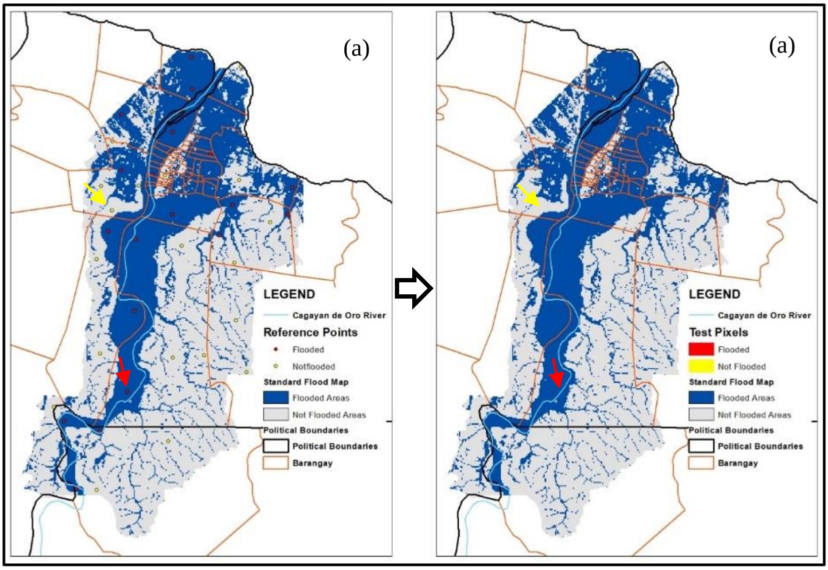

Figure 1. Creation of test pixels in red and yellow arrows (b)
from the reference points in (a)

To validate the image classification process, the classified image from MATLAB and the number codes for the two classes and from the test points were matched to check that each class was the same in both (i.e., 1 for flooded and 2 for not flooded) (Figure 2). The reference points from the standard flood map and the classified image from MATLAB were then combined using the combine tool of ArcGIS and generate an image with new combined raster values for the two images.

A confusion matrix using the pivot tool from ArcGIS was then generated and used for the evaluation of results. The confusion matrix described the performance of the linear binary classification using the set of test data from the actual (standard flood map) against that of the predicted (MATLAB generated image).

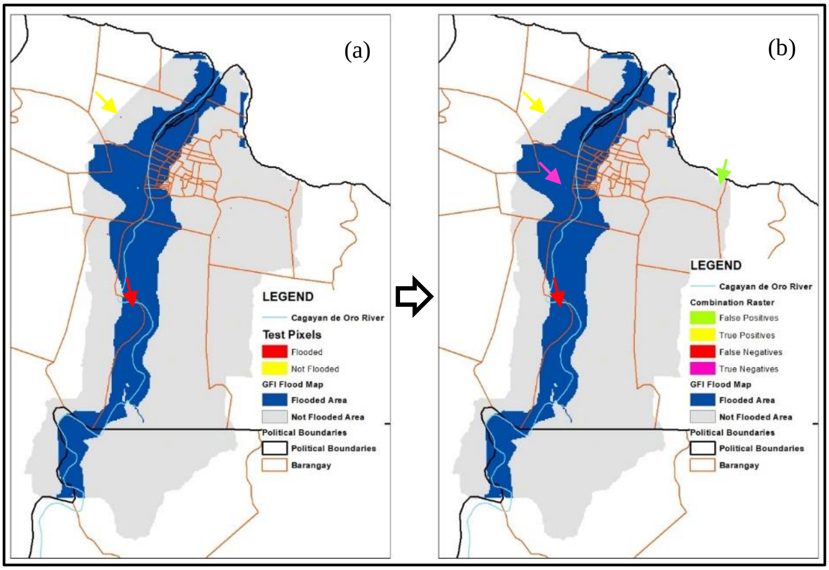

Figure 2. Matched test pixels with GFI flood map (a);
a combination of classified image (b)

2.6 Evaluation Results

To evaluate the performance of each descriptor, it was required that the number of derived test pixels be identified as True Positives (*TP*) or the correct positive prediction, False Positives (*FP*) or the incorrect positive prediction, True Negatives (*TN*) or the correct negative prediction and False Negatives (*FN*) predicted or the incorrect negative predictions based on the standard flood map. From these values, the standard measures for identifying the errors and correct predictions were the following:

a) True Positive Rate (*TPR*) or Sensitivity, which is the number of correct positive predictions among all positive samples available or

$$TPR=\frac{TP}{TP+FN} \tag{3}$$

b) False Negative Rate (*FNR*), defined as the number of incorrect negative results among the total number of positives or

$$FNR=\frac{FN}{TP+FN} \tag{4}$$

c) True Negative Rate (*TNR*) or Specificity – the number of correct negative predictions among all negative samples available or

$$TNR = \frac{TN}{TN+FP} \tag{5}$$

d) False Positive Rate (*FPR*) or (1-Specificity), defined as the number of incorrect positive predictions over the total number of negatives or

$$FPR = \frac{FP}{TN+FP} \tag{6}$$

e) Error Rate (*ERR*), defined as the number of all incorrect predictions over the total number of the dataset or

$$ERR = \frac{FP+FN}{TP+FN+TN+FP} \tag{7}$$

f) Accuracy (*ACC*), which is the number of all correct predictions over the total number of the dataset or

$$ACC = \frac{TP+TN}{TP+FN+TN+FP} \tag{8}$$

g) Kappa Coefficient (*K*), which is the measure of agreement between classification and truth values or

$$K = \frac{TOTAL\ ACCURACY\text{-}RANDOM\ ACCURACY}{1\text{-}RANDOM\ ACCURACY} \tag{9}$$

For the error rate and accuracy, the best error rate is 0.0 and the worst is 1.0 while the best accuracy is 1.0 and the worst is 0.0. The performance of the classifiers was evaluated using the Receiver Operating Characteristics (ROC) curve and the Area Under the ROC Curve (AUC) (Fawcett, 2006). The value of the AUC ranges from 0.5 (completely random classifier) to 1.0 (perfectly discriminating classifier) (Samela *et al.*, 2017).

3. Results and Discussion

3.1 Standard Flood Map versus GFI Image

The study made two major considerations in the analysis of the resulting image from the linear binary classification carried out in MATLAB. The standard flood map used in the study was generated using a 1m x 1m resolution

Lidar image, which is required in hydraulic and hydrodynamic models to provide a more accurate representation of the topography and thus allow for the accurate hydraulic simulation of flood models. For the linear binary classification, the study used a 30m x 30m resolution image. This huge difference in resolution consequently made an obvious difference in generating a more accurate output from the linear binary classification process.

Figure 3 shows the comparison between the standard flood map and GFI flood-prone areas. Given the image resolution difference used in the generation of the two maps, the map generated by the GFI was able to satisfactorily identify flooded areas along the main channel of Cagayan de Oro River. While the standard flood map shows flooding in smaller streams, these areas are historically not identified as flood-prone areas.

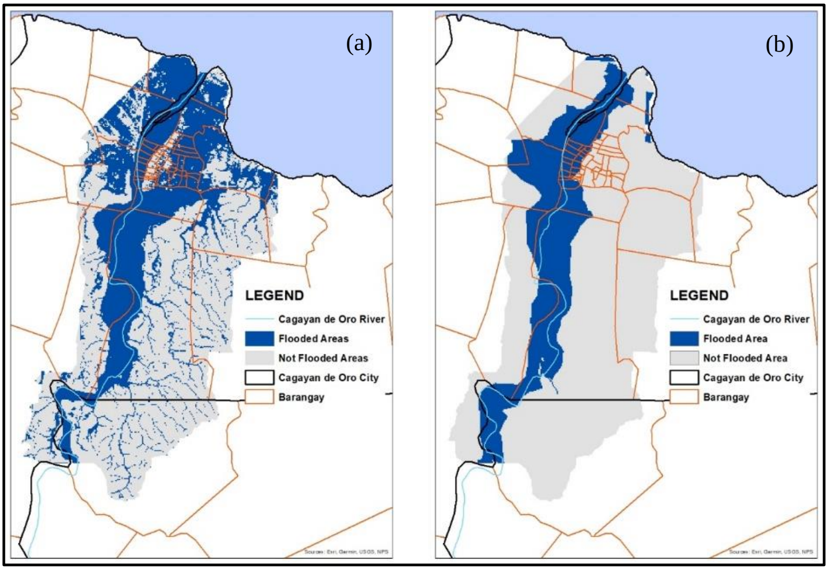

Figure 3. Standard flood map (a) vs. GFI flood-prone areas (b)

3.2 Confusion Matrix

Table 1 shows the confusion matrix generated from the image classification. Based on the confusion matrix, the study recognized from the reference points that were established, that out of the 20 test sites only 12 were accurately classified as flooded areas and eight were incorrectly classified as not flooded.

For not flooded areas, the matrix shows that out of the twenty test sites, only 17 were correctly classified as not flooded while 3 are incorrectly classified as flooded. From the confusion matrix, the study was able to calculate for the True Positive Rate or Sensitivity, True Negative Rate or Specificity, Error Rate, Accuracy, Precision, Matthew's Correlation Coefficient and the AUC.

Table 1. Confusion matrix for accuracy assessment of flood-prone areas

		Actual		
	Classification	Flooded	Not Flooded	Ground Truth
Predicted	Flooded	(*TP*) 12	(*FN*) 3	15
	Not Flooded	(*FP*) 8	(*TN*) 17	25
	Total	20	20	40

Using the confusion matrix in Table 1 and the formulas given above, values for the different metrics were calculated. The sensitivity value of 0.80 suggests that during the conduct of the accuracy assessment for flooded areas, there is an 80% probability that the test pixels identified the areas as positive or flooded areas. In the same manner, the specificity value of 0.68 indicates that during the accuracy assessment for not flooded areas, test pixels identified these areas as negative or not flooded. Table 2 below shows the summary of the calculated values of the metrics for errors and correct predictions.

Table 2. Summary of computed values for the metrics of errors and correct predictions

Measure	Calculated Value
True Positive Rate or Sensitivity (*TPR)*	0.80
True Negative Rate or Specificity (*TNR*)	0.68
False Positive Rate or 1-Specificity (*FPR*)	0.32
False Negative Rate (*FNR*)	0.20
Error Rate (*ERR*)	0.275
Accuracy (*ACC*)	0.725
Kappa Coefficient *(K*)	0.60

From Table 2, a false positive rate of 0.32 implies that there is a 32% probability that the areas identified as flooded will be identified as not flooded. Likewise, a false negative rate of 0.20 implies that 20% of the areas identified as not flooded will be identified as flooded areas. The accuracy value of 0.725 or 72.5% represents the proportion of true positive results (true positive and true negative). It indicates that 72.5% of times the test result is accurate regardless if positive or negative. While this explanation remains correct most of the time, it is of value to mention that even if the accuracy value is high, it does not directly follow that both sensitivity and specificity would also have high values (Zhu *et al.*, 2010). In terms of measuring the extent to which the test sites in the accuracy assessment were correct representations of the variables measured, a kappa coefficient of 0.60 was achieved. Table 3 is the interpretation of the kappa coefficient (McHugh, 2012), implying that the accuracy assessment has a moderate level of agreement.

Table 3. Kappa coefficient interpretation

Kappa Value	Level of Agreement	Percent Data Reliability
0 – 0.20	None	0 – 4%
0.21 – 0.39	Minimal	4 – 15%
0.40 – 0.59	Weak	15 – 35%
0.60 – 0.79	Moderate	35 – 63%
0.80 – 0.90	Strong	64 – 81%
Above 0.90	Almost Perfect	82 – 100%

Based on the test's true positive rate or sensitivity against the false positive rate or 1-specificity, a ROC curve was generated by varying the threshold from (0,0) to (1,1) and tracing the curve through the ROC space. According to Fawcett (2003), classifiers appearing on the left-hand side of the ROC curve can be thought of as "conservative" as they make positive classification only with strong evidence so they make false positive errors but have low true positive rates. On the other hand, classifiers on the upper right-hand side of the curve are thought of as "liberal", making positive classifications with weak evidence, making them classify nearly all positives correctly but often have high false positive rates.

Based on the ROC curve in Figure 4, the classifiers for this study, point (0.32, 0.8) is to perform well in the "conservative" side. The ROC point (0.32, 0.8) also has the highest accuracy of 72.5% compared to the others with accuracies ranging from 50 to 70%, which can indicate that the classifier is capable of identifying likely positives than at identifying likely negatives

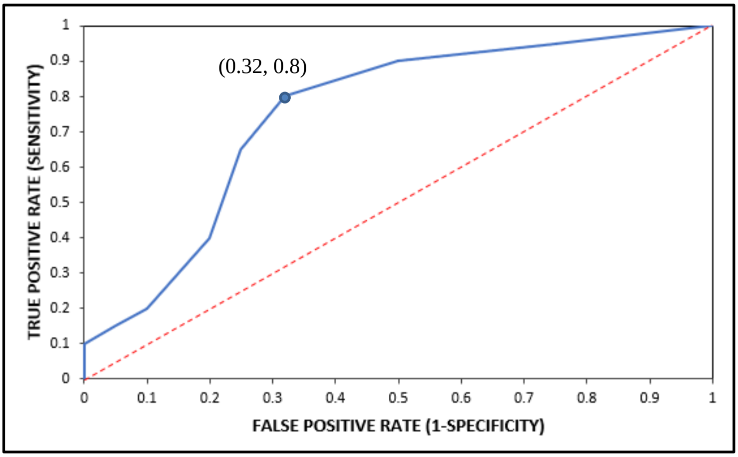

Figure 4. Receiver Operating Characteristics (ROC) Curve

To be able to summarize the performance of the classifier into a single index, an AUC can be calculated. The area under the ROC curve is a measure of how well a parameter can distinguish between two diagnostic groups, in this case, flooded and not flooded. The General Internal Medicine of the University of Nebraska Medical Center (Tape, n.d.) has outlined the rating for AUC as shown in Table 4. From the ROC curve, the computed AUC is 0.75, indicating that the parameter was able to fairly distinguish which areas are flooded and not flooded.

Table 4. Rating for Area Under the ROC Curve

Area Under the ROC Curve (AUC)	Rating
0.90 – 1.0	Excellent
0.80 – 0.90	Good
0.70 – 0.80	Fair
0.60 – 0.70	Poor
0.50 – 0.60	Fail

The AUC measures discrimination, denoting that the test is able to correctly classify those areas that are flooded and not flooded. By randomly selecting an area that is flooded and an area that is not flooded and subjecting both of them to the test, the AUC is that percentage when the test correctly classifies the areas as flooded and not flooded from the randomly selected pairs.

Overall, the accuracy assessment performed in the research study was able to provide noteworthy information as to the suitability of the linear binary classification method in distinguishing flooded and not flooded areas. The accuracy value revealed how often the classifier is correct – for this study, 72.5%. The kappa coefficient tells how well the classifier performed than it would have performed purely by chance, which is 0.60 for this study – indicating that it has data reliability of 35-65%. An AUC of 0.75 further indicates that in a randomly selected case, a flooded area has a score larger than that of a not flooded area in 75% of the time.

4. Conclusion and Recommendation

The results show that a readily accessible digital elevation model with the use of the linear binary classifier GFI allows the delineation of flood-prone areas in one of the watersheds of the Cagayan de Oro River Basin. The classifier provides acceptable detection accuracies using only a small number of datasets, very low cost and a reduced amount of data processing. Comparing the standard flood map and the GFI predicted map, flood area prediction of the GFI shows higher accuracies within the river channel as the source of hazard. This is being so as the watershed’s contributing area and drainage networks are taken as primary features for the linear binary classification. Based on the study result, prediction accuracies can be greatly influenced by the quality of the input datasets, particularly the DEMs being used. High-resolution DEMs can significantly enhance the detection accuracies of the classification method. Having shown acceptable results, the GFI method can be applied to other non-modeled or ungauged river basins although further investigation may be carried out to strengthen confidence on the use of the method.

5. Acknowledgement

The authors would like to extend their sincerest gratitude to Dr. Caterina Samela and Dr. Raffaele Albano of the Universita Degli Studi Della Basilicata, Italy for their kind assistance and for sharing their expertise and knowledge in the use of the GFI.

6. References

Bloschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: a review. Hydrological Processes, 9, 251-290.

Fawcett, T. (2006). An introduction to roc analysis. Pattern Recognition Letters, 27, 861-874.

Fawcett, T. (2003). ROC graphs: notes and practical considerations for data mining researchers. Retrieved from https://www.hpl.hp.com/techreports/2003/HPL-2003-4.pdf

Herold, C., & Mouton, F. (2011). Global flood hazard mapping using statistical peak flow estimates. Hydrology and Earth System Sciences, 8, 305-363.

Manfreda, S., Nardi, F., Samela, C., Grimaldi, S., Taramasso, A., Roth, G., & Sole, A. (2014). Investigation on the use of geomorphic approaches for the delineation of flood prone areas. Journal of Hydrology, 517, 863-876.

Manfreda, S., Samela, C., Sole, A., & Fiorentino, M. (2014). Flood-prone areas assessment using linear binary classification based on morphological indices. In: Beer, M., Au, Siu-Kui, Hall, J.W. (Eds.), Proceedings of the Second International Conference on Vulnerability and Risk Analysis and Management (ICVRAM) and the Sixth International Symposium on Uncertainty, Modeling, and Analysis (ISUMA), The University of Liverpool, The United Kingdom, 2002-2011.

McHugh, M.L. (2012). Interrater reliability: the kappa statistics. Biochemia Medica, 22(3), 276-282.

Merz, R., & Bloschl, G. (2004). Regionalisation of catchment model parameters. Journal of Hydrology, 287, 95-123.

Padi, P. T., Baldassarre G. D., & Castellarin, A. (2011). Floodplain management in africa: large scale analysis of flood data. Physics and Chemistry of the Earth, 36, 292-298.

Pappenberger, F., Dutra, E., Wetterhall, F., & Cloke, H.L. (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade. Hydrology and Earth System Sciences, 16, 4143-4156.

Samela, C., Troy, T., & Manfreda, S. (2017). Geomorphic classifiers for flood-prone areas delineation for data-scare environments. Advances in Water Resources, 102, 13-28. doi: 10.1016/j.advwatres.2017.01.007

Tape, T.G. (n.d.). Interpreting Diagnostic Tests. Retrieved the General Internal Medicine, University of Nebraska Medical Center. Retrieved from http://gim.unmc.edu/dxtests/roc3.htm

Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. Health Care and Life Sciences, NESUG Proceedings 2010, 1-9.