

## ***Interactive comment on “A Bayesian Framework Based on Gaussian Mixture Model and Radial Basis Function Fisher Discriminant Analysis for Flood Spatial Prediction (BayGmmKda V1.1)” by Dieu Tien Bui and Nhat-Duc Hoang***

**Dieu Tien Bui and Nhat-Duc Hoang**

hoangnhatduc@dtu.edu.vn

Received and published: 15 March 2017

Authors' reply to Reviewer 2's comments

Journal Name: Geosci. Model Dev

Manuscript: Bayesian Framework Based on Gaussian Mixture Model and Radial Basis Function Fisher Discriminant Analysis for Flood Spatial Prediction (BayGmmKda V1.1)

Reply to Reviewer's Comments

General comment

C1

Overall the presented work is technically interesting and contains novelties. However, major revisions are required to make it suitable for publication in GMD.

Response to reviewer's comment: We thank the reviewer for taking times and expertise to constructively comment on our manuscript. We believe that the manuscript is a meaning full contribution to the body of knowledge because this is the first time the BayGmmKda model is proposed for flood study with very promising result. We have carefully revised and provided a substantial improvement for the revised manuscript according to the reviewer's suggestions.

Regarding your comment "The paper currently does not clarify the benefits of the proposed data-intensive model for management purposes.", we 'd like to reply as follows: The model proposed in this study can make inference on the spatial prediction of flood or flood susceptibility. The current model cannot provide estimation of the magnitude and severity the events. Nevertheless, the flood susceptibility assessment is of great usefulness for local authorities in landuse planning and management by overlaying the flood susceptibility map onto planed land use maps in different scenarios. Accordingly, areas with very high flood susceptibility could be determined. Of cause, the current flood susceptibility map is only in district scale; therefore, after determining high flood susceptibility areas, larger scale studies should be carried out focusing on these areas. Based on the reviewer's comment, we will better address the benefits of the proposed model for land-use management purposes in the introduction part of the revise version as follows: "The model proposed in this study is capable of predicting the spatial prediction of flood and deriving the flood susceptibility. The derived flood susceptibility is of great usefulness for local authorities in landuse planning and management by overlaying the flood susceptibility map onto planed land use maps in different scenarios."

Regarding your comment "The presented results seem promising in the region of study but general statements about superiority of the proposed model in comparison with other techniques could only be made through evaluation in other flood prone areas.", we 'd like to reply as follows: We agree with the reviewer on this comment. The model

C2

proposed in this study will be applied for spatial modeling of flood in other study areas in Vietnam as well as in other countries. However, the data collection and processing are time-consuming. Thus, we consider this comment of the reviewer as a future research direction. We will modify the conclusion in the revised version to address this comment of the reviewer. Regarding the comment "In terms of presentation and English writing, the paper is quite poor in its current form and does not seem suitable for publication without major edition.", We will carefully check and improve the English writing of the revised manuscript.

#### Specific comments

1. Abstract is too short and not informative.

Response to reviewer's comment: We agree with the reviewer, therefore we have rewritten the abstract in the revised manuscript as follows: "Abstract. In this study, a probabilistic model, named as BayGmmKda, is proposed for flood assessment with a study area in Central Vietnam. The new model is essentially a Bayesian framework constructed a combination of Gaussian Mixture Model (GMM), Radial Basis Function Fisher Discriminant Analysis (RBFDA), and a Geographic Information System database. In order to compute the posterior probability of flood, the GMM algorithm is utilized for modeling the data distribution. Additionally, the RBFDA method is integrated into BayGmmKda to construct a latent variable that maximizes the data discrimination with respect to the two class labels of 'flood' and 'no-flood'. Experiments used for measuring the model performance point out that the hybrid framework is superior to other benchmark models including the adaptive neuro fuzzy inference system and the support vector machine. To facility the model implementation, a software program of BayGmmKda has been developed in Matlab environment. The new BayGmmKda program can accurately establish a flood susceptibility map for the study region. Accordingly, local authorities can overlay this susceptibility map onto various land-use maps for the purpose of planning or management. "

C3

2. The literature review of flood forecasting is poor. Current literature review is only focused on specific studies similar to the current work while ignoring the overall picture of flood and streamflow forecasting. Response to reviewer's comment: We understand that the overall picture of flood and streamflow forecasting is very large for the objective of this particular study. After performing literature review, we see that there are two groups of approaches for flood and streamflow forecasting: (i) the first one is "regression modeling" and (ii) the second one is "classification modeling" The first approach group has used for very long time, but required detailed monitoring data for modeling, these data are difficult to obtain for Vietnam as a developing country. The modeling result of this approach group could provide spatial and temporal prediction of flood for study areas. The second group is relatively new and does not require flood monitoring data. It uses "on – flood pixel" and "off - non flood pixel" for flood modeling. Therefore, this approach is feasible for modeling of large areas with the use of remote sensing and GIS data. In other words, the input-output datasets in this approach is very different with those of the traditional approaches. The modeling result of this approach group provide only where flood may occur (spatial prediction of flood or flood susceptibility), this does not provide temporal prediction or flood discharge. In this study, we use the second approach group, therefore we have mainly specific studies similar to our works and we think that the current literature review is reasonable.

3. What is the definition of flood used in this study? What is the difference between flood and no-flood? How severe an event needs to be to be called flood? How can predictions be useful for government agencies without providing an estimation of the magnitude and severity of the events? Response to reviewer's comment: Floods in this study are flood locations that occurred in the study areas and have been determined based on documentary sources of the local district, interpretation of Landsat 8 Operational Land Imagery. In addition, flood locations were collected during field works using handheld GPS. Non-flood points were randomly generated from non-flood areas within the study area based on DEM, i.e. ridges (we has used DEM to generate topographical shades i.e. flat, Ridge, Saddle Ravine, Convex hillside, Saddle hillside, Slope hillside,

C4

Concave hillside, Inflection hillside). Regarding your comment “How severe an event needs to be to be called flood?”, we’d like to reply as follows: In this study, only flash flood is modelled. The flood locations used in this study were provided by the local authority. With the current database, the exact information on the severity of these floods is not available. However, all 76 floods in this study caused huge damages to the local people. Regarding the comment, “How can predictions be useful for government agencies without providing an estimation of the magnitude and severity of the events?” ,we’d like to response as follows: As we explained in the above answer. The flood model in this study could provide only where flood may occur called spatial prediction of flood or flood susceptibility. The current model is not able to deliver estimation of the magnitude and severity the events. The flood susceptibility is still useful for local authorities in landuse planning and management by overlaying the flood susceptibility map to planed land use maps in different scenario. Subsequently, areas with very high flood susceptibility could be determined. Of course, the flood susceptibility map is only in district scale; therefore, after determining high flood susceptibility areas, larger scale studies should be carried out focusing on these areas. 4. Flood points are used in this study, and not flood areas, with GIS maps. So for information on flood influencing factors in each of these flood points, how many map pixels were used? Was each flood point only associated with the pixel it was located in? If more map pixels than one were used to get information on flood influencing factors for each flood point, how was the area of analysis (relevant pixels) determined for each flood point?

Response to reviewer’s comment: We have 76 flood polygons in the flood inventory map. However, we used 76 points for these polygons and these points were determined based on overlaying these polygons and the DEM. The grid (20 m size) of the district map was constructed by 3900 columns and 4125 rows.

Regarding the question “Was each flood point only associated with the pixel it was located in?. If more map pixels than one were used to get information on flood influencing factors for each flood point, how was the area of analysis (relevant pixels)

C5

determined for each flood point?”: Yes, each flood point only associated with the pixel it was located in. No more map pixels than one were used to get information on flood influencing factors for each flood point.

Technical corrections There are too many instances of poor English writing throughout the paper to be listed. Major edition seems necessary to make the paper suitable for publication.

Response to reviewer’s comment: According to the reviewer’s comment, the whole manuscript will be proofread to improve the English writing.

---

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-311, 2017.

C6