Natural Hazards
and Earth System
Sciences
Open Access
Discussions

# Flood hazard in the Mekong Delta – a probabilistic, bivariate, and non-stationary analysis with a short-termed future perspective

N. V. Dung[1,3], B. Merz[1], A. Bárdossy[2], and H. Apel[1]

[1]GFZ German Research Centre for Geosciences, Section 5.4 Hydrology, Potsdam, Germany
[2]Institute of Hydraulic Engineering, University of Stuttgart, Stuttgart 70569, Germany
[3]Southern Institute of Water Resources Research SIWRR, Ho Chi Minh City, Vietnam

Correspondence to: N. V. Dung (dung-viet.nguyen@gfz-potsdam.de)

## Abstract

In this paper we present a novel approach for flood hazard analysis of the whole Mekong Delta with a particular focus on the Vietnamese part. Based on previous studies identifying the flood regime in the Mekong delta as non-stationary (Delgado et al., 2010), we develop a non-stationary approach for flood hazard analysis. Moreover, the approach is also bi-variate, as the flood severity in the Mekong Delta is determined by both maximum discharge and flood volume, which determines the flood duration. Probabilities of occurrences of peak discharge and flood volume are estimated by a copula. The flood discharges and volumes are used to derive synthetic hydrographs, which in turn constitute the upper boundary condition for a large-scale hydrodynamic model covering the whole Mekong Delta. The hydrodynamic model transforms the hydrographs into hazard maps. In addition, we extrapolate the observed trends in flood peak and volume and their associated non-stationary extreme value distributions to the year 2030 in order to give a flood hazard estimate for the near future. The uncertainty of extreme flood events in terms of different possible combinations of peak discharge and flood volume given by the copula is considered. Also, the uncertainty in flood hydrograph shape is combined with parameter uncertainty of the hydrodynamic model in a Monte Carlo framework yielding uncertainty estimates in terms of quantile flood maps. The proposed methodology sets the frame for the development of probabilistic flood hazard maps for the entire Mekong Delta. The combination of bi-variate, non-stationary extreme value statistics with large-scale flood inundation modeling and uncertainty quantification is novel in itself. Moreover, it is in particular novel for the Mekong Delta: a region where not even a standard hazard analysis based on a univariate, stationary extreme value statistic exists.

# 1 Introduction

Flooding, defined as destructive abundance of water, is one of the most damaging natural disasters. However, flooding in some regions in the world (e.g. the Mekong Delta) has to be considered in a wider perspective. Floods can also be beneficial by, for example, bringing sediments and nutrients to fields of flooded regions where agriculture is the main economic activity. In fact, floods have been one important factor in the development of modern civilization, as the settlements of early civilization along the Nile and the Euphrat show. Recently, owning to climate change and other factors like demographic and economic developments, it is widely acknowledged that an increasing number of people are threatened by floods, especially in coastal and estuarine regions (e.g. Balica et al., 2012; Bhuiyan et al., 2012; Merz et al., 2010; IPPC, 2012). Assessing flood risk, preparing effective flood mitigation measures and utilizing flood benefits at the same time have thus become an even more vital task in water resources planning and flood management.

Floods in the Mekong Delta occur annually with a trend of increasing inter-annual variability (Delgado et al., 2010). Average floods are perceived as beneficial to the Delta. The annual floods are the basis of the livelihoods of several million people in the Cambodian and Vietnamese part of the Delta. People. The ecosystem and agriculture are well adapted to and benefit from the occurrence of floods. However, extreme floods (e.g. flood years 1961, 1971, 2000, 2001 and 2002) can cause huge damage (Hoa et al., 2007; MRC, 2009) and pose a serious threat to millions of people. This flood hazard is likely to increase due to observed and expected sea level rise (Doyle et al., 2010; Hoa et al., 2007; Wassmann et al., 2004; MONRE, 2009; IPCC, 2007). Changes in the Mekong flood hydrology are however uncertain, as GCM predictions of future climates show a high uncertainty in the predictions of future flood discharge for the Mekong (Kingston et al., 2011; compared to Hoanh, 2010; MONRE, 2009). Contrary to Kingston et al. (2011) the latter two studies are based on a single GCM projection only, thus the findings and conclusions have to be treated with caution and serve as a

proof for the uncertain predictions when compared with each other and with Kingston et al. (2011). But in any case climate change poses an additional challenge to flood management. Summarizing the above mentioned facts, it is rather surprising that a flood hazard or even a risk assessment for the Mekong Delta does not exist, neither in published literature, nor, to the authors' knowledge, within Vietnamese offices and authorities.

Flood management should be based on a sound hazard analysis (Merz et al., 2010). Extreme value statistics for hydrological data, e.g. precipitation or flow discharges, typically form its basis, taking stationarity and independence of the data series as underlying assumptions. However, these assumptions remain doubtful because of the change in natural variability, especially in the context of climate change (Khaliq et al., 2006; Adlouni et al., 2007; Strupczewski et al., 2001). For the Mekong basin, Delgado et al. (2010) identified the flood regime as significantly non-stationary and advocated the use of non-stationary methods for flood hazard estimation in the Mekong. Thus, in this study, beside the use of the traditional approach of flood frequency analysis based on the stationary assumption, a new approach taking the non-stationarity into account is developed. This approach is adopted not only to the estimations of the probabilities of occurrence of annual maximum discharge as a proxy for flood severity, but also to the flood volume as another very important flood feature in the Mekong Delta. The dependence and joint probabilities of peak discharge and volume are modeled by copulas.

In order to generate synthetic but representative flood hydrographs for Kratie – the gauge representing the upper boundary of the Mekong Delta – the method of Apel et al. (2004) was adopted to extract typical standardized flood hydrographs from the discharge time series by cluster analysis. These typical hydrographs have an associated empirical probability of occurrence and are scaled by the peak flow and volume pairs taken from the copula. The combination of different pairs of peak flow and volume having the same probability of occurrence with different typical synthetic hydrographs allows for an estimation of the uncertainty inherent in the synthetic flood events for a given probability of occurrence.

The other key method in flood hazard analysis is numerical simulation of inundation processes for the estimation of inundated areas and depths (Apel et al., 2004; Vorogushyn et al., 2010). For the Mekong Delta Dung et al. (2011) developed a large-scale quasi-2-D hydraulic model covering the complete domain. The model was automatically calibrated with a multi-objective optimization algorithm against in-situ data from a network of gauging stations and against a series of inundation maps derived from ENVISAT Advanced Synthetic Aperture Radar (ASAR) satellite images. The calibration method yielded a number of Pareto-optimal parameterizations, which represent the parameter uncertainty inherent in the hydraulic simulations.

The non-stationary bivariate frequency analysis and the hydraulic model developed by Dung et al. (2011) are combined to generate probabilistic flood hazard maps for floods of different magnitudes. The probabilistic maps indicate the uncertainty inherent in the hydrological input and in the hydraulic simulation, which is quantified by a Monte Carlo procedure. The uncertainty is presented in quantile maps (5 %, 50 %, 95 %) of maximum inundation depth and inundation duration, following recent research in flood hazard analysis (Vorogushyn et al., 2010; Domeneghetti et al., 2012). The presented approach is the first published flood hazard analysis for the Mekong Delta. Moreover, the overall concept of a bivariate and non-stationary frequency analysis combined with a large-scale inundation model is novel in flood hazard analysis in general. Similar studies are not known to the authors.

The paper is organized as follows: Sect. 2 describes the Mekong Delta and data used, followed by a description of the methods in Sects. 3–5. A particular focus is put on the bivariate non-stationary frequency analysis. In Sect. 6 the derivation of the probabilistic flood hazard maps is described and the results and uncertainties are discussed in comparison to a standard stationary approach. In addition, a short projection of the future flood hazard is performed assuming that the observed trends remain stable in the next two decades.

**Flood hazard in the Mekong Delta**

N. V. Dung et al.

◄◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## 2 Data and methods

### 2.1 The Mekong Delta

The Mekong River Delta is situated in the Cambodian and Vietnamese part of the lower Mekong basin (Fig. 1). The flow regime in the delta is highly complex due to the very low topography of the Mekong Delta, the complex channel system especially in the Vietnamese part with multiple layered circle channels, two different ocean tidal modes, the peculiarity of the upstream boundary in Cambodia with the Great Lake Tonle Sap and the numerous man-made hydraulic structures in the Vietnamese part of the Delta. The most striking natural feature of the delta is the Tonle Sap lake in Cambodia, which serves as a huge buffer for flood discharges during the monsoon season. When the water level in Kratie, which is the upper boundary of the Delta, exceeds the level of 17 m a.s.l., overbank flow in the area of Kampong Cham downstream of Kratie is initiated (Hoi, 2005). The overbank flow separates into two different flow paths: The right overbank flow follows the main stream, but divides further near Phnom Penh into flow into the Tonle Sap Lake to the North and further down in southern direction to the Mekong Delta, Vietnam. A large proportion of the left overbank flow flows directly south and reaches the Plain of Reeds, the North-eastern part of the Delta in Vietnam, causing a second flood wave besides the floods from the Mekong main stem in these parts.

The floodplains in Cambodia differ significantly from those in Vietnam. In Cambodia the plains are in a comparable natural state with only a few control structures, whereas the floodplains in Vietnam are under large regulation by a huge system of navigation and irrigation channels, sluice gates, pumps and especially an extensive dike system. Thus, the combination of the natural hydraulic peculiarities in combination with the large anthropogenic influence creates a very complex system, which challenges every modeling ambition.

## 2.2 The flood inundation model

The flood inundation model was set up to represent the river network and floodplains in the Mekong Delta. The model domain having the size of more than $55\,000\,\text{km}^2$ embraces the complete Delta from Kratie including the Tonle Sap Great Lake in Cambodia to the river mouths in Vietnam (Fig. 2). For such a large-scale model only a 1-D approach is feasible from a computational point of view. However, the model needs to represent the floodplains in order to be hydraulically meaningful and to enable prediction of inundation extent. Two different ways of representing the floodplains are followed. The comparatively natural floodplains in Cambodia including the Tonle Sap, where hardly any channels and dikes exist, are simulated by wide cross sections including the floodplains, which is the usual approach in 1-D hydrodynamic flood models. The floodplains in Vietnam are treated separately. They are separated into a multitude of compartments enclosed by high dikes, comparable to dike rings in the Netherlands. Because most of the compartments represent a closed system surrounded by dikes and channels, floodplain compartments ("flood cells") are modeled by artificial branches with low and wide cross sections extracted directly from the Shuttle Radar Topography Mission (SRTM) DEM with a horizontal resolution of 90 m. Those branches are linked to the channels by control structures. Weirs were used to represent dike overflow, and sluice gates were implemented in the model whenever information on existing sluice gates was present. Figure 2 illustrates the approach, by which a quasi-2-D model is established. This approach resembles the first hydrodynamic model developed for the Mekong Delta by Cunge (1975) but adopted to the present channels and dikes, which were hardly present at that time. The model consists of 4235 branches, of which more than 2134 are used to represent the flood plains of the Mekong Delta in Vietnam in 564 compartments, which is equivalent to about 26 376 computational nodes. The length of the simulated channel system is about 25 000 km. The topographical data for the model were collected from various sources, thus having different levels of accuracy. The most accurate data could be collected for the main stems of the Mekong and some larger

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

channels. For some smaller channels, where no bathymetric data were available, bed elevation had to be assumed. The same holds true for the dike elevations, with the additional problem of different datums used by the different agencies and districts for geo-referencing of elevations.

5    The flood model was calibrated automatically for the flood year 2008 and validated for 2009 (Dung, 2011a; Dung et al., 2011b). The calibration and validation were seen from a multi-objective point of view using both high-temporal, low-spatial resolution data (gauge hydrographs) and low-temporal, high-spatial resolution data (inundation areas from remote sensing). The first objective function (hereafter $F_1$) was formulated

10  using the most commonly used measure in the calibration of hydrological models: the Nash-Sutcliffe coefficient of efficiency. The second objective function (hereafter $F_2$) was defined utilizing the flood-area index, which compares binary patterns (flooded, not flooded). The goal of the calibration was to find the best parameter sets representing both objectives. For more details, interested readers are referred to Dung et al. (2011b).

15  The calibration process was successfully achieved and the resulting Pareto-front of the final calibration set is shown in Fig. 3. The final front exhibits a large spread showing trade-off solutions. It means that an improved performance in one of the two objectives, which measure the model skill with respect to the temporal or spatial aspects, can only be achieved at the expense of the other objective. This also means in consequence

20  that model parameterizations providing satisfying performance in both temporal and spatial dynamics cannot be found with the given model setup. Thus no general recommendable parameter set exists and an appropriate parameterization has to be chosen according to the specific purpose of different model applications.

In order to account for this model/parameter uncertainty, we selected two parameter

25  sets in this study. We do not use all Pareto optimal sets, because of computational time constraints when performing a full Monte Carlo uncertainty analysis over all parameter sets. The uncertainty of the hydrodynamic model is thus not considered in a full probabilistic way, but in scenarios. The parameter sets selected are: the best with respect to $F_2$, because spatial extent has a higher relevance for flood hazard in the Delta, and

**Flood hazard in the Mekong Delta**

N. V. Dung et al.

Printer-friendly Version

Interactive Discussion

the best Euclidean distance to the optimal solution (hereafter EU, for more information see Dung et al., 2011). The EU parameter set "harmonizes" the model performance of the two objectives.

Overall, the validation yielded results comparable to the calibration as Fig. 3 (right) illustrates. Here the performance measures of calibration and validation are plotted against each other. For both $F_1$ and $F_2$ the scatter plots show an acceptable performance over all pareto optimal parameter sets (lower right row), with a slightly lower performance in the intermediate range between the optimal parameters sets for $F_1$ and $F_2$.

## 2.3 Discharge series at Kratie

Kratie in Cambodia represents the upper boundary of the Mekong Delta, respectively the transition from the Lower Mekong Basin to the Delta. A daily discharge series from 1 January 1924 to 31 December 2009 is used for the flood hazard analysis. Discharge at Kratie is considered as the drainage of the Mekong basin area of about 646 000 km$^2$. The maximum specific runoff (unit discharge) is reported as one of the biggest in the world. Its value is about 0.12 m$^3$/s/km$^2$ (O'Connor and Costa, 2004; MRC, 2005). For each year a fixed flood period from the beginning of June to the end of November is selected for the hazard analysis. This period covers the annual flood season of the Mekong (MRC, 2005), i.e. the peak of each flood season always falls in this period. The flood volume is consequently calculated as the flood volume within this fixed period. The volume of flow may be expressed in cubic meters but this leads to very large numbers and therefore large volume units are commonly used (Reddy, 2005). In this study, the volume is calculated using the unit of day.m$^3$/s (or cumec-day = 86 400 m$^3$ s$^{-1}$) instead of m$^3$. Figure 4 (left) shows the annual peak flow, the average flow and the minimum flow series of the flood period and the corresponding annual flood volume series. Figure 4 (right) illustrates that the peak flow and volume are roughly linearly correlated. Table 1 lists some basic statistic parameters of the peak and volume series. Both series are slightly skewed to the right. The peak series shows stronger skewness,

while the skewness in the volume series is very small. The Pearson correlation coefficient is 0.73, which implies a quite large linear dependence.

## 2.4 The flood hazard analysis framework

In order to derive synthetic discharge hydrographs at Kratie, the three essential variables flood peak, flood volume and hydrograph shape are analyzed. This is a necessary extension to the standard flood hazard analysis procedures based on annual maximum flows. In the Mekong Delta not only flood peak discharge is determining the flood hazard, but also the flood volume, as the devastating flood in the year 2000 illustrates. In this year the peak discharge was not extraordinarily high, but the flood volume was the largest recorded causing deep inundation through prolonged water logging and ponding in the flat floodplains of the Delta (Hoi, 2005; Hoc, 2000; MRC, 2005). Therefore we developed a copula-based analysis dealing with the bivariate frequency analysis of both flood peak flow and volume. A cluster analysis is applied to identify characteristic hydrograph shapes along with their empirical probability of occurrence. The characteristic flood hydrographs are scaled by discharge and volume; their probabilities of occurrence, resp. return interval are taken from the bivariate frequency distribution. Finally, flood hazard maps for the whole Mekong Delta are derived by the calibrated large-scale hydrodynamic model of the Mekong Delta driven by the hydrographs derived for Kratie.

To account for uncertainties in the hazard assessment, a Latin Hypercube – Monte Carlo framework sampling from combinations of flood peak and volume with identical probabilities of occurrences and different hydrograph shapes is applied. Figure 5 illustrates the general procedure of the synthetic hydrograph generation for the hazard analysis. By this method probabilistic hazard maps are produced for selected probabilities of occurrences. Additionally, we apply both stationary and non-stationary frequency distributions and compare the results in terms of hazard maps derived by the different approaches. This comparison allows to discuss the errors in the hazard analysis

introduced by neglecting the reported non-stationarity in flood variability (Delgado et al., 2010).

## 3  Bivariate flood frequency analysis via copulas

Classical bivariate approaches are only applicable when the margins of the two random variables (or their transformations) follow the same family of distributions (e.g Yue, 1999, 2000, 2001; Yue et al., 1999; Adamson et al., 1999). Based on the information gained from an uncorrelated univariate approach (Dung, 2011), a classical bivariate log-normal distribution could be theoretically applied (Hiemstra et al., 1976). However, a more flexible approach based on copulas is adopted for this study in order to retain generality and transferability of the approach. Copulas enable the use of different probability distribution functions for the different variables to be described, e.g. a generalized extreme value distribution for peak series and Gamma distribution for volume series.

The copula concept emerged a long time ago when Sklar's theorem (Joe, 1997; Nelsen, 2005) was first stated in 1959. Recently, copula-based techniques have been frequently applied to a wide range of applications (Genest and Favre, 2007), because copula-based models are able to overcome the noted limitations of classical multivariate analysis. While copulas have been applied very often in "high-risk" financial and accrual sectors, its use in water resources is quite recent, for example for analyzing rainfall data (Singh et al., 2005), flood behavior (Salvadori and De Michele, 2004; Klein et al., 2010) or geostatistical groundwater quality models (Bárdossy, 2006).

### 3.1  Basic concept of copulas

A very short introduction to the concepts of copulas is given. Details can be found in several text books (Joe, 1997; Nelsen, 2005; Cherubini et al., 2004; Salvadori et al., 2005).

A multivariate distribution with support on the $n$-dimensional unit hypercube and uniform margins is called a copula. Copulas are bounded by the Frechet-Hoeffding bounds. Given marginals, it is always possible to gain an infinite number of multivariate joint distributions. The main advantage of copulas is their "scale invariant" property, meaning that they remain the same by any scaling transformation. Another advantage of copulas is the separation of dependence function and marginal; hence, the dependence of the random variables is not influenced by their marginals.

For the presented case of flood peak discharge and volume at Kratie different copulas are tested. Firstly, the Gaussian copula is selected because both marginal ($q$ and $v$) were well fitted to the log-normal distribution which originates from the normal distribution family. Additionally, one-parameter Archimedean copulas are applied. Archimedean copulas are among the most frequently used, especially in hydrology (Genest and Favre, 2007). From the Archimedean copula family, the Ali-Mikhail-Haq copula is not considered because its application is based on the cases where the rank dependence is rather small (<0.4), which is not the case in the presented study. Hence, the four candidate bivariate copulas Gumbel-Houggard, Frank, Clayton and Gaussian described in Table 2 are further investigated.

When applying copulas for extreme value statistics, tail dependence is of high importance. Tail dependence is referred to the "amount of dependence" in the upper right quadrant tail (or lower left quadrant) of a bivariate distribution. More precise definition and discussion are given in Joe (1997) and Salvadori et al. (2005). A copula can have upper tail dependence or lower tail dependence or both. And it can also have no tail dependence at all. Among the candidate copulas, Gaussian is a symmetric copula without tail dependency. The Frank copula does not have tail dependency either. Gumbel-Houggard has upper tail dependency and the Clayton copula has lower tail dependency. Hence, care should be taken when selecting a copula family for frequency analysis of extreme value statistics.

## 3.2 Parameter estimation and copula identification

When a parametric copula is selected to model the dependence between two random variables, parameters related to that copula family must be estimated. Different estimation methods such as direct method, pseudo maximum likelihood (PML) or canonical maximum likelihood (CML) method, inference from margin method (IFM) and exact maximum likelihood (EML) exist (Joe, 2005; Genest and Favre, 2007; Cherubini et al., 2004). We use IFM for estimating the copula parameters for both stationary and non-stationary analysis. For the evaluation of the fitting performance the Akaike information criteria AIC (Akaike, 1974) is used. The AIC is a way of selecting a model from a set of models under consideration. The AIC criterion adopted in this case is defined as:

$$\text{AIC} = -2 \times \text{loglikelihoood} + 2 \times \text{number of parameters to be estimated.} \qquad (5)$$

In general, the best model has the minimum AIC value. However, in practice, when examining the performance of a model in a set of candidate models, it is necessary to examine how much its AIC is larger than the minimum AIC of the model set. Guidelines for applying the AIC are given in Brunham and Anderson (2002).

### 3.2.1 Estimation results

Table 3 (first line) shows the parameters and corresponding AIC values for the estimation of the four candidate copulas for the stationary analysis. The Gaussian copula outperforms the other candidates with respect to the AIC value. Figure 6 (left) provides a graphical test of the goodness-of-fit. A random sample of size 20 000 (small dots) from each of the four copulas with parameters estimated by the IFM method was generated. The two blue lines represent the threshold value corresponding to the 100-yr return period estimated by the three-parameter log-normal distributions (Hosking, 1993) fitted to peak and volume time series independently. The 86 pairs of peak and volume observations are represented by black circles. The plot allows judging the viability of the copula model. The two copulas Gumbel and Frank fit well to the observed

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

data. However, the Gaussian copula seems to be the best (within the range of the observed data), while the Clayton copula is obviously not a suitable choice as it spreads out significantly in the upper part. And most of the points do not drop in the upper right quadrant formed by the two blue lines. Because we target estimating floods with a return interval of up to 100 yr, the performance of the copulas up to this probability is of superior importance for the selection of the copula. Therefore, together with the fact that it has the best AIC value, the Gaussian copula is selected.

### 3.2.2 Non-stationary analysis

Delgado et al. (2010) detected trends in the annual maximum discharge time series of the Mekong at Kratie (falling trend in mean, rising trend in variance) and stated that the stationarity assumption does not hold. They advocate the use of non-stationary frequency distributions to estimate flood probabilities. We build on this finding, test for trends in the volume and discharge series and find suitable non-stationary distributions for both peak and volume. Tests of different non-stationary distribution functions identified the three-parameter lognormal distributions as the most appropriate choice (Dung, 2011). As illustrated in Fig. 7, both volume and peak discharge series show a negative trend in the location parameter (= falling mean) and a positive trend in the scale parameter (= rising variance). These trends will be extrapolated to estimate the flood hazard in "a near" future. The readers who are interested in the non-stationary univariate analysis of the two time series are referred to Dung (2011).

Using the identified distributions, the suitability of the different copulas is tested. Again, in terms of the AIC value, the Gaussian copula shows the best fit (Table 3). The Clayton and the Frank copula are eliminated because of their significantly higher AIC. For the graphical check of the suitability a small modification was made. In the non-stationary analysis, data is time-dependent. Therefore, standardization by "removing time" from the data is needed. In Fig. 8, this standardization does not change the rank of the data when the stationary model is applied (left) but does change when the non-stationary model is applied (right) which implies a change in the dependence structure

of the data (rank correlation) leading to the change in the copula based model. The standardization procedure is described in detail in Hosking and Wallis (1997).

We assume that the copula for the original series and the copula for the transformed series remain the same. Based on this assumption, the graphical validity check is given in Fig. 6 (right). The same argument as in the stationary analysis holds. Again, we target at estimating floods with a return interval of 100 yr. Hence, the Gaussian copula is also the best choice for the non-stationary analysis.

### 3.2.3 Return period via copula

The return period of an event is the reciprocal probability of occurrence. Following (Salvadori et al., 2005, 2010, 2011), in a bivariate frequency analysis the most common definition of return periods is given by the average lag years between two flood events of which:

– either volume ($V$) or peak ($Q$) exceed certain values ("OR" case return period),

– both volume and peak exceed certain values ("AND" case return period).

Given a pair ($qv$) of a random bivariate ($QV$), the "OR"- return period, $T_{QV}(q,v)$, is formulated as:

$$T_{QV}(q,v) = \frac{1}{\Pr(Q \geq q\, \text{OR}\, V \geq v)} = \frac{1}{1 - C(1 - \frac{1}{T_Q(q)}, 1 - \frac{1}{T_V(v)})} \tag{6}$$

The "AND" – return period, $T^{QV}(q,v)$, is formulated as:

$$T^{QV}(q,v) = \frac{1}{\Pr(Q \geq q\, \text{AND}\, V \geq v)} = \frac{1}{\frac{1}{T_Q(q)} + \frac{1}{T_V(v)} - \frac{1}{T_{QV}(q,v)}} \tag{7}$$

where Pr(−) denotes the probability of an outcome, $C(-)$ the copula function, $T_X(x)$ univariate return period of the random variable $X$ at the given value $x$ ($X$ standing for $Q$ or $V$, $x$ standing for $q$ or $v$).

From Eqs. (6) and (7) it is obvious that if the return periods of the two marginal distributions $T_Q T_V$ are known and the copula function is defined, the corresponding bivariate return period can be derived. However, when the return periods of the bivariate analysis and the copula function are known, infinite pairs of ($T_Q T_V$) can be derived. Furthermore,

$$T_{QV}(q,v) \leq T_Q(q) T_V(v) \leq T^{QV}(q,v) \tag{8}$$

i.e. the return period of a pair of ($qv$) given by the "OR" case is smaller or equal to the individual univariate return periods (= has higher or equal probability), which in turn have lower return periods than the "AND" case. This means in turn, when performing a hazard analysis for defined return periods, peak discharge and flood volume fulfilling the "AND" case must be smaller compared to the "OR" case for every distinct return period.

Which copula return period is chosen ("AND" or "OR") cannot be defined a priori and has to be decided case wise dependent on the observed data and purpose of the study. To this end, we plotted the observed peak and volume pairs along with the different return periods $Q-V$-pairs given by "AND" and "OR" definition in Fig. 9. It can be seen that the observed time period is better covered by the "AND" definition of the return periods in both the stationary and non-stationary analysis. The different return period lines separate the observed data reasonably well, whereas with the "OR" definition most of the observed data fall below the 10-yr return period, even the event of 2000 with the historically largest damage. This is not plausible, and thus the "AND" definition is selected.

## 4  Hydrograph shape analysis

To provide the complete hydrograph for given pairs of flood peak and volume, the method of Apel et al. (2004, 2006) is used. It is based on standardization and clustering of observed flood hydrographs. First, hydrographs are normalized to the peak flow, i.e. dividing the series by the peak value of each event. Next, a cluster analysis classifies the normalized hydrographs into groups based on their similarity. A simple hierarchical approach, the ward method, is applied. The similarity between hydrographs is determined by the squared Euclidean distance.

For the hydrographs at Kratie, four clusters are chosen. This choice is – to a certain extent – subjective. The number of hydrograph clusters should be large enough to depict the variability of the hydrograph shape but small enough in order to identify characteristic hydrographs, but also to alleviate the problem of computational demand when dealing with a Monte Carlo (or Latin Hypercube) simulation.

The dendrogram in Fig. 10 shows the result of the cluster analysis. The y-axis shows the "distance" assigned to all intermediate groups of normalized hydrograph shapes. The shaded yellow region gives the threshold for the four cluster selection. To each cluster an occurrence probability is assigned, given as the number of events in each cluster divided by the total number of hydrographs (see the right plot of Fig. 10).

Figure 11 (left) shows the four non-dimensional characteristic hydrographs computed as the mean hydrograph from all events within each cluster. The four hydrographs can be broadly classified as single peaked flood events, distinguished in broad central peak, narrow central peak, early, and late peak. Figure 11 (right) shows that no correlation between the occurrence of a certain hydrograph class and flood magnitude and volume exists. Only the narrow central peak hydrographs seem to have a tendency towards smaller flood volumes. Hence, the approach to derive a pair of flood peak and volume, and to cluster the hydrograph shapes independently of peak and volume is valid.

# 5   Generation of 100-yr flood hydrographs

Synthetic hydrographs are generated using the characteristic hydrographs scaled up by peak discharge and flood volume, whereas both are estimated from the stationary and non-stationary bivariate distributions.

As mentioned, we aim at quantifying the flood hazard of 100-yr event floods examplarily. Two reference years are selected for the probabilistic flood hazard mapping. The year 2009 is used as baseline scenario or current status, and 2030 is defined as the future time horizon. In order to estimate the flood hazard in 2030, the detected trends in the location and shape parameters of the non-stationary distributions are extrapolated to the year 2030. We assume hereby that the detected trends will continue in the near future, which is justifiable from our point of view given the observed monotonous trends. However, further extrapolation to a more distant time horizon should be avoided due to the unquantifiable uncertainty associated.

Figure 12 illustrates the three level curves corresponding to 100-yr event. $Q - V$-pairs along these curves are the bivariate events with the same associated return period. However, the coincidence of $Q$ and $V$ along these curves are not equally likely, i.e. a very high flood peak is unlikely to occur along with a very low volume. Combinations along the 100-yr return period curves can differ greatly in terms of their values. For example, moving along the solid red curve, the 3 $(qv)$ points of $(33\,273\,\text{m}^3\,\text{s}^{-1},\ 5\,706\,900\,\text{day.m}^3/\text{s})$, $(63\,223\,\text{m}^3\,\text{s}^{-1},\ 5\,432\,700\,\text{day.m}^3/\text{s})$ – the red vertex, $(66\,588\,\text{m}^3\,\text{s}^{-1},\ 2\,437\,700\,\text{day.m}^3/\text{s})$ have the same return period value of 100 yr, but their probability of occurrences (density) along the curve are $2.5 \times 10^{-11}$, $1.36 \times 10^{-4}$, $4.4 \times 10^{-13}$, respectively. The vertex of the intersection between the 100-yr bivariate return period curve and the curve describing the full dependence (see Chebana et al., 2011; Volpi et al., 2012) between peak and volume has the highest density along the bi-variate return period curve. In other words, this is the most likely bi-variate 100 yr flood event.

For the probabilistic hazard analysis a Monte Carlo sampling was used to generate 100 pairs of ($qv$) for the stationary and non-stationary analysis, respectively. By this technique the 100-yr event is simulated in a probabilistic framework considering the uncertainty of the bi-variate 100-yr event from the sampling of the distribution functions and the dependency between flood peak and flood volume. The sampled pairs of $q$ and $v$ were associated with a randomly selected normalized characteristic hydrograph conditioned by their probability of occurrences. Finally, synthetic hydrographs are generated by rescaling the normalized characteristic hydrographs first to sampled $q$ and consequently to the sampled volume $v$ while retaining $q$.

## 6   Derivation of probabilistic flood hazard maps

For the reference years 2009 and 2030, probabilistic flood hazard maps for the 100-yr event are derived by the stationary and non-stationary approach, respectively. The synthetic hydrographs are transferred into inundation areas utilizing the calibrated and validated hydrodynamic model described in Sect. 2.2. In order to account for model parameter uncertainty, we use those roughness parameter sets from the final pareto-optimal parameter sets giving the best result in terms of simulation of inundation extend (best $F_2$) and the one with shortest Euclidean distance to the optimal simulation. The simulation period covers the flood period from 1 July to 30 November. As the actual operation of the sluice gates is not known, we assume that all the sluice gates in the high dike compartments remain closed, i.e. three crops are grown per year. This represents a scenario aiming at a maximization of economic profit. Maximum water levels at more than 26 000 computational nodes are interpolated to 2-D maps with horizontal resolution of 90 m considering the closed high dike compartments. In addition, maps of inundation duration and water depth are derived. These maps are interpreted in a probabilistic way by calculating inundation statistics for every grid cell. Commonly used percentiles, 5 %, 50 % (median), 95 %, are utilized to illustrate the aleatory and epistemic uncertainty in the flood hazard analysis.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

# 7 Results and discussion

Figures 13 and 14 show the flood hazard as maps of the 100-yr event in terms of maximum inundation depth using either the best Euclidean distance parameter set for the hydraulic model, or the best parameter set for the simulation of inundation extent ($F_2$). Figures 15 and 16 present inundation duration derived by the same two hydraulic parameter sets. In every figure the different rows provide the results from a stationary analysis based on the observed time series until 2009 (top row), the results from the non-stationary analysis for the reference year 2009 based on the observed time series (middle row), and the projected hazard for 2030 derived from the extrapolation of the observed trends in mean and variance (bottom row). The columns of the figures show the uncertainty stemming from the aleatory uncertainty of different $Q - V$-pairs quantified by 5%, 50% and 95% pixel-wise quantiles of the set of inundation maps. The middle column presents the median in absolute values. The 5% and 95% quantile maps do not show the absolute value, but the difference to the median maps. We choose this presentation in inter-quantile maps to enhance the interpretation of the uncertainty in these large-scale maps.

It can be seen that uncertainty in the flood frequency analysis mainly affects the Cambodian part of the Delta and the northern part in Vietnam (e.g. Plain of Reeds). These regions have high inundation depths, ranging from 3 m to more than 6 m and may be inundated up to 6 months. The regions that are close to the Vietnamese coast, especially the Western coast, are less flooded and the hazard analysis is also less uncertain in this region. However, it has to be noted that the presented study does not consider climate change impacts on sea level. For the oceans around the Mekong Delta a significant sea level rise is predicted (Doyle et al., 2010; IPCC, 2007; MONRE, 2009), which in turn increases the flood hazard in the near-coastal regions. The quantification of this impact is, however, out of scope of this study.

Another interesting feature is observed mainly in the North-Western area of the Vietnamese Delta (Long-Xuyen quadrangle) and the area between the main branches of

the Mekong in the Vietnamese Delta (Hau and Tien river), but also in some parts of the Plain of Reeds: There are areas which are not inundated in most of the maps shown. These areas are protected by closed ring dikes, which are in most cases designed to protect against a flood as experienced in the year 2000. In the presented simulations

5   we assume that the sluice gates in these high dikes remain closed during the flood season top grow a third crop per year. The event in 2000 approximates a 100-yr event, both for the stationary and non-stationary statistical analysis. Thus it is interesting to observe that some protected areas (e.g. those are illustrated by the yellow circles in the top-right of Fig. 14) are flooded in the 95 % quantile maps indicating a remaining

10  probability of flooding even for events having a comparable likelihood to the one they were designed to withstand. But as Dung et al. (2011) by the automatic calibration of the model have shown, this could also be an artifact of erroneous dike elevation defined in the hydrodynamic model.

Generally the differences in maximum inundation depth and inundation duration be-

15  tween the quantile maps are much larger than the differences between the stationary (for the year 2009) and non-stationary (for 2009 and 2030) approaches. This means that the uncertainty due to likely temporal changes of the hazard plays a comparatively small role. The projection of future flood hazard in 2030 by trend extrapolation shows slightly reduced median maximum inundation depths. This effect is more pronounced

20  with the $F_2$ parameter set, but also smaller than the inter-quantile differences. The projection of the inundation duration shows hardly any change compared to the recent situation.

A closer look at the differences between the EU and $F_2$ parameter sets, i.e. the uncertainty in hydrodynamic modeling, shows that the $F_2$ parameter set yields higher

25  inundation depths and longer durations in the Vietnamese part of the Delta. The opposite effect can be observed for the Cambodian part. This is a direct consequence of the model parameterization and calibration. The automatic calibration routine altered the hydraulics in the Cambodian part in order to best represent the inundation extent in the Vietnamese part (Dung et al., 2011). But overall, the differences between the

two model parameter sets are small compared to the inter-quantile differences. Hence, the parameter uncertainty of the hydraulic model has a rather small influence. However, these findings may change when floods of different magnitudes are considered. For example, for more frequent floods, the differences between the two model parameterizations should become more obvious, as the inundation depth and extent in the Vietnamese part of the Delta are controlled by the dike elevations. The presented 100-yr events overflow most of the dikes in any case, thus yielding only small differences in the hazard maps. For smaller floods the parameterization should be more influential, as the $F_2$ parameter sets creates higher water levels in the rivers in the Vietnamese part of the delta and can thus overflow some dikes already, whereas the EU set may produce water levels below most of the dike lines. Dung et al. (2011b) calibrated and validated the hydrodynamic model for the years 2008 and 2009. Both of them were average floods. Here significant differences in inundation extent and depth between the two parameter sets could be observed, and we expect that this would also be the case in the hazard assessment of frequent floods. Presenting these maps would however exceed the scope of a journal publication.

The uncertainty associated with the temporal change should increase for more extreme floods. This follows from the larger differences in the tails of the stationary and non-stationary distributions. This has direct implications for the usefulness of the presented approach: while we can conclude that for the assessment of the hazard by events up to the 100-yr flood, a standard stationary approach would be acceptable from a practical point of view, even for the estimation of the hazard in the near future, this statement cannot be supported if a full hazard assessment including extreme events is conducted. In this case the use of the non-stationary approach has to be regarded as more appropriate.

While the copula concept in general proved to be an appropriate choice for the studied problem, it has to be noted that for the estimation of events with return periods higher than the presented 100-yr event, the tail dependency in the bivariate analysis becomes more important. The selected Gaussian copula does not take tail dependency

**Flood hazard in the Mekong Delta**

N. V. Dung et al.

into account. For the analysis of extreme events the tail dependency of peak discharge and volume has to be studied in more detail. If tail dependency proves to be important, a Gumbel or $t$-copula would be possible choices. However, as the length of the observed time seriesords is limited (84 yr), this investigation is difficult to perform. The uncertainty associated to the estimation of extreme events is dominated by the length of the data series, just as in a univariate analysis (see e.g. Apel et al., 2008), and not by the copula used. For this reason we restricted the analysis up to 100-yr events for which the Gaussian copula is an appropriate choice.

## 8   Conclusions

In this study a comprehensive approach for flood hazard mapping at large scales is presented and exemplarily applied to the Mekong Delta. The approach links a multivariate, non-stationary flood frequency analysis with a large-scale hydraulic model for simulating the inundation in spatial context. It also takes into account the main uncertainties, namely the uncertainty in the hydrological input to the Mekong Delta as well as in the hydraulic model parameterization. It results in probabilistic flood inundation maps for defined probabilities of occurrence, from which we present hazard maps for 100-yr events.

Bivariate flood frequency analysis and cluster analysis are applied to study the important hydrograph characteristics: peak discharge, flood volume and hydrograph shape. A method based on copulas is used to construct a bivariate flood peak and volume distribution and to analyze the bivariate frequency. This is an essential component for flood hazard estimation in the Mekong Delta as both components define flood severity, either by themselves or in conjunction. Thus the multivariate statistical analysis has large advantages compared to a standard univariate approach. In fact, a flood hazard analysis for the Mekong Delta has to be multivariate in order to be meaningful. Besides the traditional approach of flood frequency analysis based on the stationarity assumption, an approach taking non-stationarity into account was also examined.

The analysis has shown that both flood peak and flood volume can be modeled as log-normal distributions in a stationary and non-stationary analysis. Negative trends in the location parameter and a positive trend in the scale parameter were detected in both peak discharge and volume time series. From the results it can be concluded that for estimating the 100-yr event a stationary analysis is practically acceptable, although theoretically not sound. Nevertheless, we advocate for the use of the non-stationary approach as it is (a) theoretically sound, (b) able to provide short-term projections of future flood hazard, and (c) more likely to adequately capture the hazard, if we include extreme events much larger than the 100-yr flood in the analysis. The last point, however, is only inferred theoretically and not proven by results.

The analysis of the different uncertainty sources – hydrological input and parameter uncertainty of the hydrodynamic model – revealed that the uncertainty in the hydrological input has the largest impact on the inundation maps. It has to be noted that uncertainty in hydrological input refers to the possible combinations of peak discharge and flood volume with identical joint probabilities of occurrence, not to parameter uncertainty of the fit of the distribution functions to the data. Hence, this dominant source of uncertainty is aleatory not epistemic uncertainty (Merz and Thieken, 2005). This uncertainty can by definition not be reduced by better techniques or data, as it depicts the natural variability. Should the observed rising trend of this variability continue to increase in the coming decades, increasing uncertainty has to be expected.

Flood risk management in the Mekong Delta should take this fact into account and add flexible measures to the usual procedure, which is typically raising dikes. The concept of "living with floods" appears to offer more appropriate solutions and would build on the long history of living with and from floods in the Delta. However, the society must be prepared that these floods will become more variable and potentially more extreme (in both directions: high and low).

In contrast to hydrological uncertainty the parameter uncertainty of the hydrodynamic model can be reduced, which is particularly relevant for the estimation of highly probable, i.e. frequent floods. Efforts should be made to improve the representation of the

dike lines, in particular the elevations, in the model to reconcile the spatial and temporal performance and to enable more reliable estimation of frequent floods.

It has to be noted that the statement, that the hydraulic parameter uncertainty is comparably small, is to some extent a consequence of the choice of only two parameter sets. It has to be expected that the consideration of the full range of Pareto-optimal parameter sets derived by Dung et al. (2011) in the Monte Carlo analysis would increase the uncertainty associated to the hydrodynamic modeling. On the other hand it can be argued, as in fact we do, that it may not be appropriate to use those parameter sets that underestimate the inundation extent.

In summary, we regard the presented study as an essential step towards risk-based flood management in the Mekong Delta. Not only because it is the first of its kind, but also because of the methods applied, which acknowledge the particular hydrological and hydraulic features of the region and the Delta. The work will be continued by attempts to reduce the epistemic uncertainties, but also by combining it with vulnerability aspects in order to assess also risks, not only hazards.

# References

Adamson, P. T., Metcalfe, A. V., and Parmentier, B.: Bivariate extreme value distributions: An application of the gibbs sampler to the analysis of floods, Water Resour. Res., 35, 2825–2832, 1999.

Adlouni, S. E., Ouarda, T. B. M. J., Zhang, X., Roy, R., and Bobee, B.: Generalized maximum likelihood estimators for the nonstationary generalized extreme value model, Water Resour. Res., 43, doi:10.1029/2005WR004545, 2007.

Akaike, H.: A new look at the statistical model identification, IEEE Trans. Automatic Control, 19, 716–723, doi:10.1109/TAC.1974.1100705, 1974.

Apel, H., Thieken, A. H., Merz, B., and Blöschl, G.: Flood risk assessment and associated uncertainty, Nat. Hazards Earth Syst. Sci., 4, 295–308, doi:10.5194/nhess-4-295-2004, 2004.

Apel, H., Thieken, A. H., Merz, B., and Blöschl, G.: A probabilistic modelling system for assessing flood risks, Nat. Hazards, 38, 79–100, 2006.

Apel, H., Aronica, G., Kreibich, H., and Thieken, A.: Flood risk analyses – how detailed do we need to be?, Nat. Hazards, 49, 79–98, 2009.

Balica, S. F., Wright, N. G., van der Meulen, F.: A flood vulnerability index for coastal cities and its use in assessing climate change impacts, Nat. Hazards, 64, 73–105, 2012.

Bárdossy, A.: Copula-based geostatistical models for groundwater quality parameters, Water Resour. Res., 42, 1–12, doi:10.1029/2005wr004754, 2006.

Bhuiyan, M. and Dutta, D.: Analysis of flood vulnerability and assessment of the impacts in coastal zones of Bangladesh due to potential sea-level rise, Nat. Hazards, 61, 729–743, 2012.

Brunham, K. P. and Anderson, D. R.: Model selection and inference: A practical information-theoretic approach, 2nd Edn., Springer-Verlag, New York, USA, 488 pp., 2002.

Chebana, F. and Ouarda, T. B. M. J.: Multivariate quantiles in hydrological frequency analysis, Environmetrics, 22, 63–78, doi:10.1002/env.1027, 2011.

Cherubini, U., Luciano, E., and Vecchiato, W.: Copula methods in finance, John Wiley & Sons, Ltd, Chichester, 308 pp., 2004.

Cunge, J. A.: Two-dimensional modeling of floodplains, Water Resour. Publications, 2, 705–762, 1975.

Delgado, J. M., Apel, H., and Merz, B.: Flood trends and variability in the Mekong river, Hydrol. Earth Syst. Sci., 14, 407–418, doi:10.5194/hess-14-407-2010, 2010.

Domeneghetti, A., Vorogushyn, S., Castellarin, A., Merz, B., and Brath, A.: Effects of rating-curve uncertainty on probabilistic flood mapping, Hydrol. Earth Syst. Sci. Discuss., 9, 9809–9845, doi:10.5194/hessd-9-9809-2012, 2012.

**Flood hazard in the Mekong Delta**

N. V. Dung et al.

Doyle, T. W., Day, R. H., and Michot, T. C.: Development of sea level rise scenarios for climate change assessments of the mekong delta, vietnam, US Geological Survey Open-File Report 2010–1165, 110, 2010.

Dung, N. V.: Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the mekong delta, PhD Thesis, University of Stuttgart, Stuttgart, available at: http://elib.uni-stuttgart.de/opus/volltexte/2012/6831/, 177 pp., 2011a.

Dung, N. V., Merz, B., Bárdossy, A., Thang, T. D., and Apel, H.: Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data, Hydrol. Earth Syst. Sci., 15, 1339–1354, doi:10.5194/hess-15-1339-2011, 2011b.

Genest, C. and Favre, A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask, J. Hydrol. Eng., 12, 347–347, doi:10.1061/(asce)1084-0699(2007)12:4(347), 2007.

Hiemstra, L. A. V., Zuchhini, W. S., and Pegram, G. G. S.: A method of finding the familiy of runhydrographs for given return periods, J. Hydrol., 30, 95–103, 1976.

Hoa, L. T. V., Nhan, N. H., Wolanski, E., Tran, T. C., and Haruyama, S.: The combined impact on the flooding in vietnam's mekong river delta of local man-made structures, sea level rise, and dams upstream in the river catchment, Estuarine, Coast. Shelf Sci., 71, 110–116, 2007.

Hoanh, C. T., Jirayoot, K., Lacombe, G., and Srinetr, V.: Impacts of climate change and development on mekong flow regime, first assessment – 2009, Mekong River Commission, Vientiane, Lao PDR, 2010.

Hoc, D. X.: Assessment on the process of the flood year 2000 and its impact on the plain of reeds, vietnam, University of Water Resources, Ho Chi Minh City, 67 pp., 2000.

Hoi, T. N.: Research on scientific and technological solutions on building dike system for sustainable development in the mekong delta, Southern Institute of Water Resources Research, Ho Chi Minh City, 2005.

Hosking, J. R. M. and Wallis, J. R.: Regional frequency analysis, Cambridge University Press, New York, 238 pp., 1997.

IPCC: Synthesis report, Contribution of working groups i, ii and iii to the fourth assessment report on the intergovernmental panel on climate change Core Writing Team, Pachauri, R. K and Reisinger, A. (Eds.), IPCC, Geneva, Switzerland, 104 pp., 2007.

IPCC: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation – Special Report of the Intergovernmental Panel on Climate Change, 2012.

Joe, H.: Multivariate models and dependence concepts, Chapman & Hall, USA, 418 pp., 1997.

Joe, H.: Asymptotic efficiency of the two-stage estimation method for copula-based models, J. Multivar. Analys., 94, 401–419, doi:10.1016/j.jmva.2004.06.003, 2005.

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J.-C., Gachon, P., and Bobee, B.: Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, J. Hydrol., 329, 534–552, 2006.

Kingston, D. G., Thompson, J. R., and Kite, G.: Uncertainty in climate change projections of discharge for the Mekong River Basin, Hydrol. Earth Syst. Sci., 15, 1459–1471, doi:10.5194/hess-15-1459-2011, 2011

Klein, B., Pahlow, M., Hundecha, Y., and Schumann, A.: Probability analysis of hydrological loads for the design of flood control systems using copulas, J. Hydrol. Eng., 15, 360–369, doi:10.1061/(asce)he.1943-5584.0000204, 2010.

Merz, B. and Thieken, B.: Separating aleatory and epistemic uncertainty in flood frequency analysis, J. Hydrol., 309, 114–132, doi:10.1016/j.jhydrol.2004.11.015, 2005.

Merz, B., Hall, J., Disse, M., and Schumann, A.: Fluvial flood risk management in a changing world, Nat. Hazards Earth Syst. Sci., 10, 509–527, doi:10.5194/nhess-10-509-2010, 2010.

MONRE: Climate change, sea level rise scenarios for vietnam, Ministry of Natural Resources and Environment, Ha Noi, 34 pp., 2009.

MRC: Overview of the hydrology of the mekong basin, Mekong River Commission, Vientiane, 73 pp., 2005.

MRC: Annual mekong flood report 2008, Mekong River Commission, Vientiane, 2009.

Nelsen, R. B.: Introduction to copulas, 2nd Edn., Springer, New York, 2005.

O'Connor, J. E. and Costa, J. E.: The world's largest floods, past and present – their causes and magnitudes, US Geological Survey Circular 1254, 13 pp., 2004.

Reddy, P. J. R.: A textbook of hydrology, Firewall Media, Delhi, India, 530 pp., 2005.

Renard, B. and Lang, M.: Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology, Adv. Water Resour., 30, 897–912, doi:10.1016/j.advwatres.2006.08.001, 2007.

Salvadori, G. and De Michele, C.: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, Water Resour. Res., 40, 1–17, 2004.

Salvadori, G. and De Michele, C.: Multivariate multiparameter extreme value models and return periods: A copula approach, Water Resour. Res., 46, W10501, doi:10.1029/2009WR009040, 2010.

**Flood hazard in the Mekong Delta**

N. V. Dung et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Salvadori, G., Michele, C. D., Kottegoda, N. T., and Rosso, R.: Extremes in nature: An approach using copulas, Springer, Dordrecht, The Netherlands, 303 pp., 2005.

Salvadori, G., De Michele, C., and Durante, F.: On the return period and design in a multivariate framework, Hydrol. Earth Syst. Sci., 15, 3293–3305, doi:10.5194/hess-15-3293-2011, 2011.

Singh, V. P., Wang, S. X., and Zhang, L.: Frequency analysis of nonidentically distributed hydrologic flood data, J. Hydrol., 307, 175–195, 2005.

Strupczewski, W., Singh, V., and Feluch, W.: Non-stationary approach to at-site flood frequency modelling i. Maximum likelihood estimation, J. Hydrol., 248, 123–142, doi:10.1016/s0022-1694(01)00397-3, 2001.

Volpi, E. and Fiori, A.: Design event selection in bivariate hydrological frequency analysis, Hydrol. Sci. J., 57, 1506–1515, doi:10.1080/02626667.2012.726357, 2012.

Vorogushyn, S., Merz, B., Lindenschmidt, K.-E., and Apel, H.: A new methodology for flood hazard assessment considering dike breaches, Water Resour. Res., 46, doi:10.1029/2009WR008475, 2010.

Wassmann, R., Hien, N. X., Hoanh, C. T., and Tuong, T. P.: Sea level rise affecting the vietnamese mekong delta: Water elevation in the flood season and implications for rice production, in: Climate change, Kluwer Academic Publishers, 89–107, 2004.

Yue, S.: Applying the bivariate normal distribution to flood frequency analysis, Water Int., 24, 248–252, 1999.

Yue, S.: The bivariate lognormal distribution to model a multivariate flood episode, Hydrol. Process., 14, 2575–2588, 2000.

Yue, S.: A bivariate gamma distribution for use in multivariate flood frequency analysis, Hydrol. Process., 15, 1033–1045, doi:10.1002/hyp.259, 2001.

Yue, S., Ouarda, T. B. M. J., Bobee, B., Legendre, P., and Bruneau, P.: The gumbel mixed model for flood frequency analysis, J. Hydrol., 226, 88–100, 1999.

**Flood hazard in the Mekong Delta**

N. V. Dung et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►◄

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Table 1.** Statistics of flood peak and volume series at Kratie.

| | Flood peak $(\text{m}^3\,\text{s}^{-1})$ | Flood volume $(10^3\,\text{day.m}^3/\text{s})$ |
|---|---|---|
| Mean | 48 529 | 4178 |
| Standard deviation | 7366 | 642 |
| *Coefficent of skewness* | *0.2782* | *0.0868* |
| *Coefficent of kurtosis* | *3.8344* | *2.9008* |
| Minimum | 29 920 | 2565 |
| Maximum | 72 141 | 5757 |
| Median | 48 309 | 4160 |
| Interquantile Range | 8684 | 818 |

**Table 2.** Summary of the four candidate bivariate copulas (Nelsen, 2005; Renard and Lang, 2007; Bárdossy, 2006).

| Name | Descriptions |
|---|---|
| Gumbel-Hougaard | $C_\theta(u,v) = \exp\left(-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^{1/\theta}\right) \quad \theta \in [1, +\infty)$ (1) |
| Frank | $C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + (e^{-\theta u} - 1)(e^{-\theta v} - 1)(e^{-\theta} - 1)\right), \quad \theta \in (-\infty, +\infty)\backslash\{0\}$ (2) |
| Clayton | $C_\theta(u,v) = \left(u^\theta + u^\theta - 1\right)^{-1/\theta}, \theta \in (0, +\infty)$ (3) |
| Gaussian | $C_\theta(u,v) = \Phi_\theta(\phi^{-1}(u) + \phi^{-1}(v)), \quad \theta \in [-1, 1]$ (4) |
| where | $(uv) \in [0,1]^2$ |
| | $\theta$ represents the copula parameter of which |
| | $C_\theta$ represents the copula function |
| | $\phi^{-1}$ represents the inverse function of the standard normal distribution cdf |
| | $\Phi_\theta$ represents the bivariate normal distribution cdf with correlation $\theta$ |

**Table 3.** Copula parameter ($\theta$) estimation for the stationary and non-stationary case.

| | Gumbel | | Frank | | Clayton | | Gaussian | |
|---|---|---|---|---|---|---|---|---|
| | $\theta$ | AIC | $\theta$ | AIC | $\theta$ | AIC | $\theta$ | AIC |
| Stationary | 1.872 | −54.1 | 6.007 | −54.2 | 1.338 | −45.9 | 0.730 | −63.6 |
| Non-stationary | 1.765 | −51.0 | 3.960 | −42.2 | 0.890 | −29.0 | 0.665 | −52.2 |
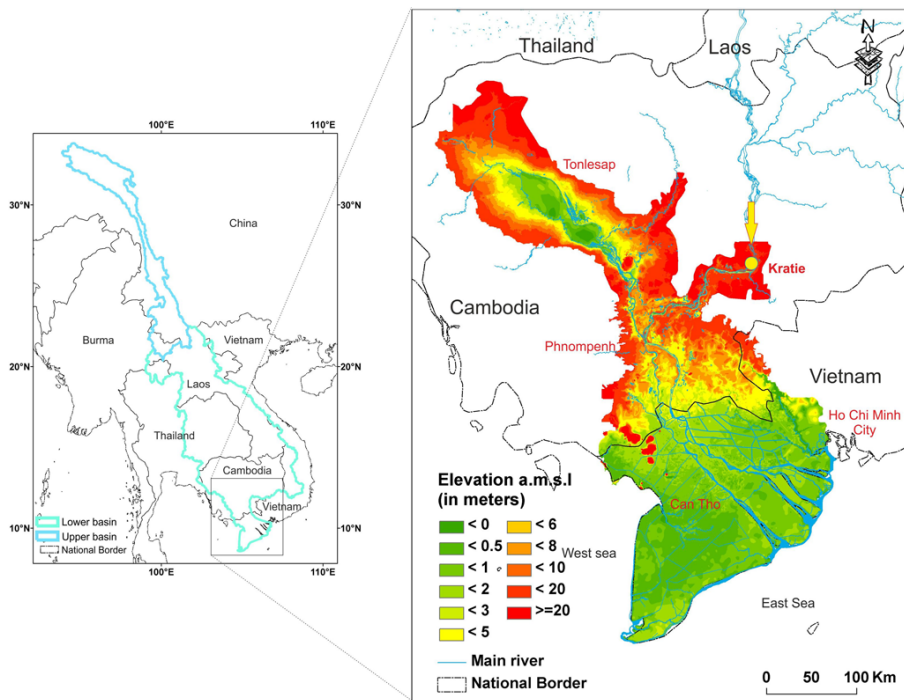
**Fig. 1.** The Mekong basin (left) and the Mekong Delta (right); the yellow filled circle with arrow indicating the main flow input to the Delta at Kratie, Cambodia.
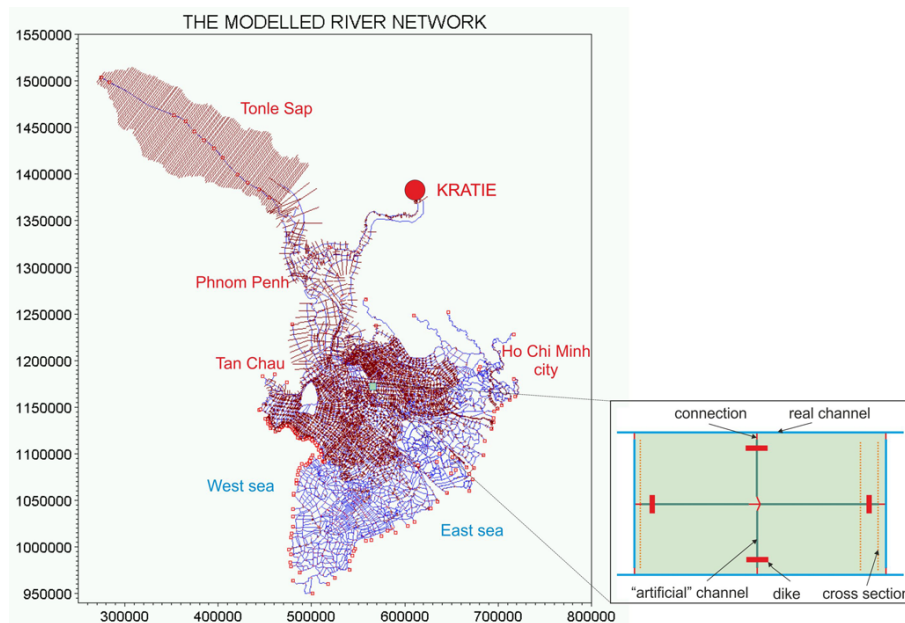
**Fig. 2.** The model river network including Tonlé Sap – Great Lake (left) and model representation of a typical floodplain in the Mekong Delta, Vietnam (right).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
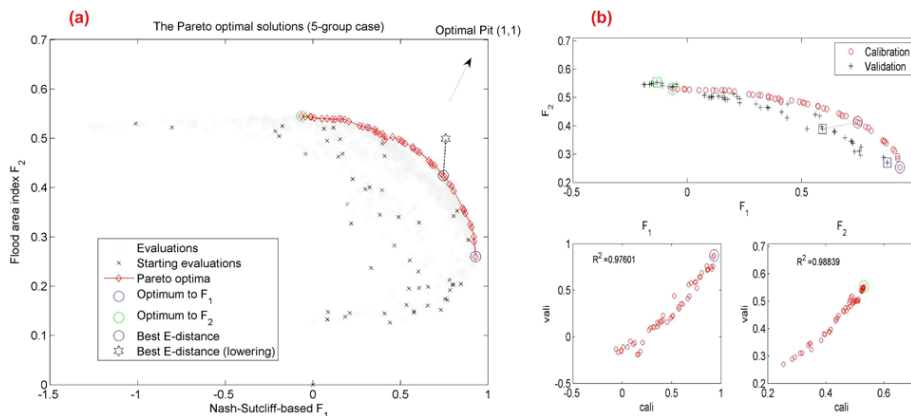
**Fig. 3.** Results of **(a)** calibration and **(b)** validation of the flood model (Calibration for flood period in 2008, validation for flood period 2009, from Dung et al., 2011 and Dung, 2011).
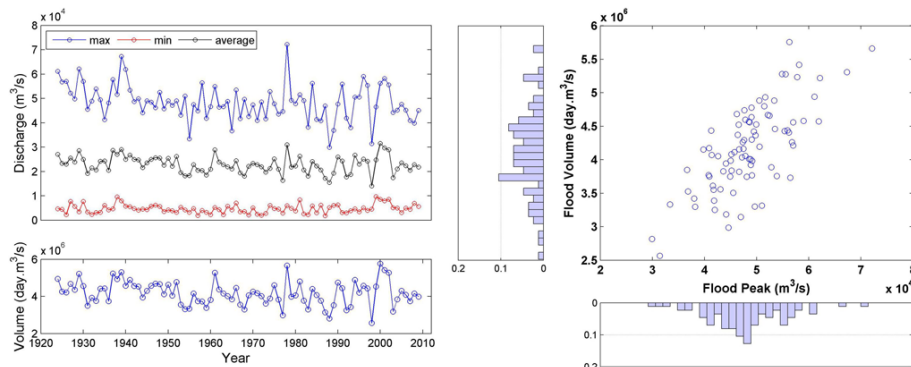
**Fig. 4.** Left: discharge and flood volume time series; Right: scatter plot of the flood volume and flood peak and histograms for 1924–2009 at Kratie.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
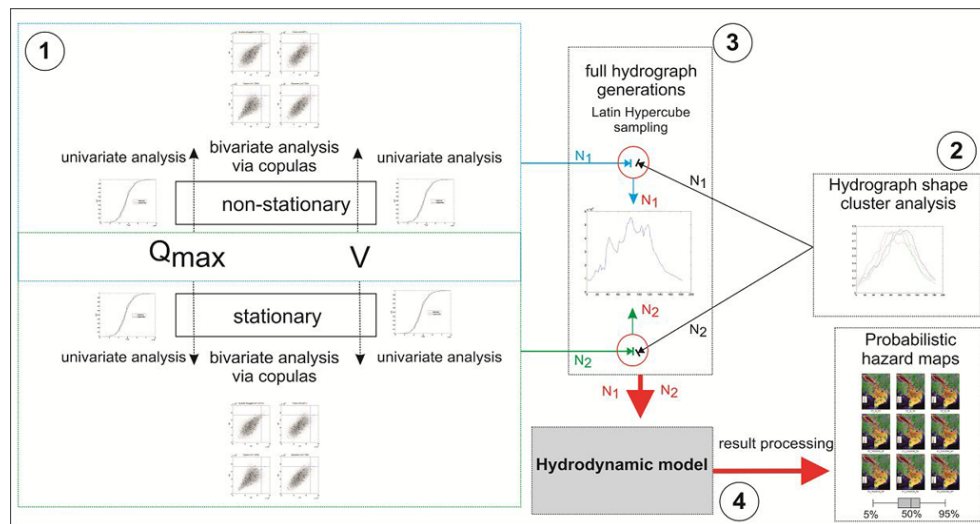


**Fig. 5.** The flood hazard analysis framework: (1) frequency analysis on the peak and volume series, (2) hydrograph shape cluster analysis, (3) generation of flood hydrographs and associated probability of occurrence and (4) flood hazard modeling and mapping with different parameter (roughness) sets.
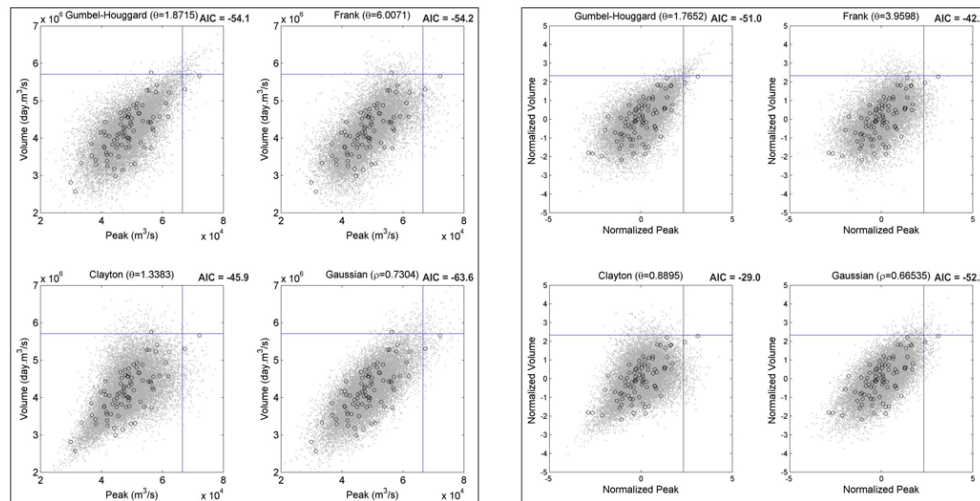
**Fig. 6.** Comparison of the observed values of flood peak and volume (circles) and 20 000 synthetic values (dots) generated via the four candidate copulas for the stationary (left) and non-stationary (right) assumption.
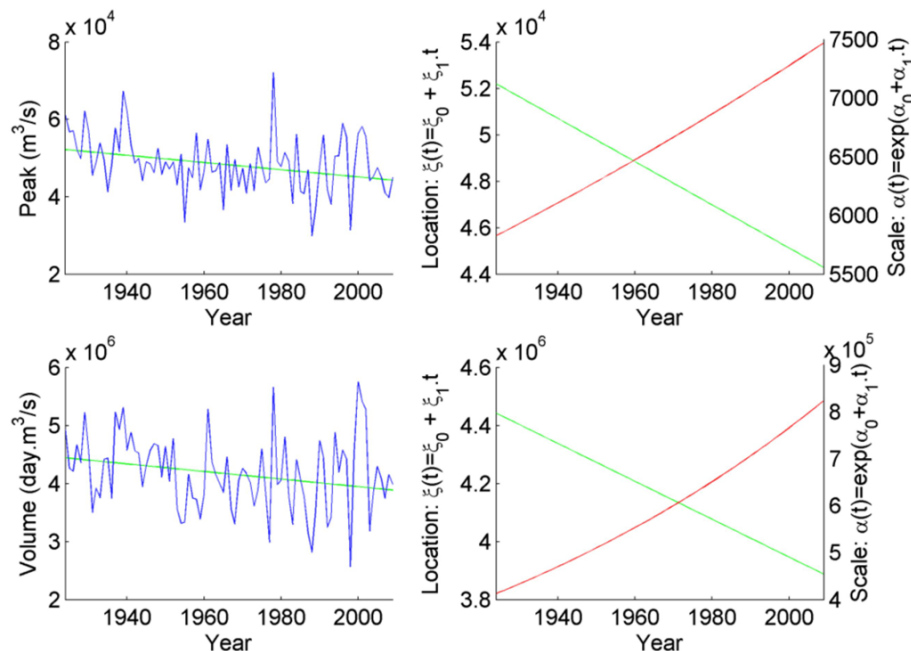
**Fig. 7.** Time varying parameters of the non-stationary models for flood peak (top) and volume (bottom), blue for the time series, green for location parameter, red for scale parameter.
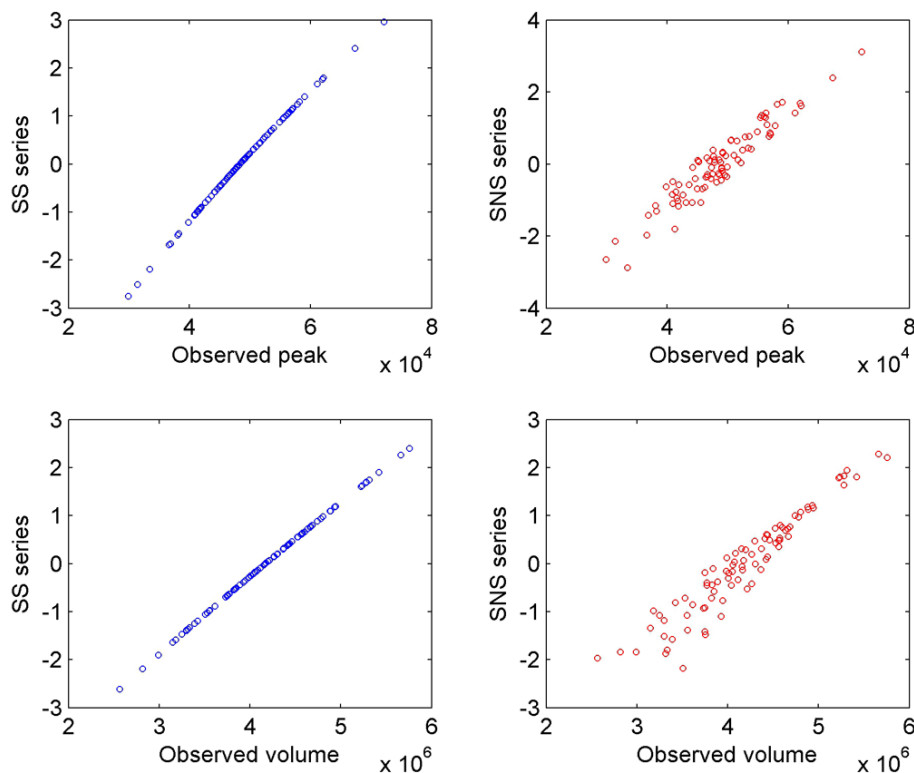
**Fig. 8.** Normalized observed peak values (top) and observed volume values (bottom) for the stationary (left) and the non-stationary analysis (right); labels of the vertical axis indicate standardized values.
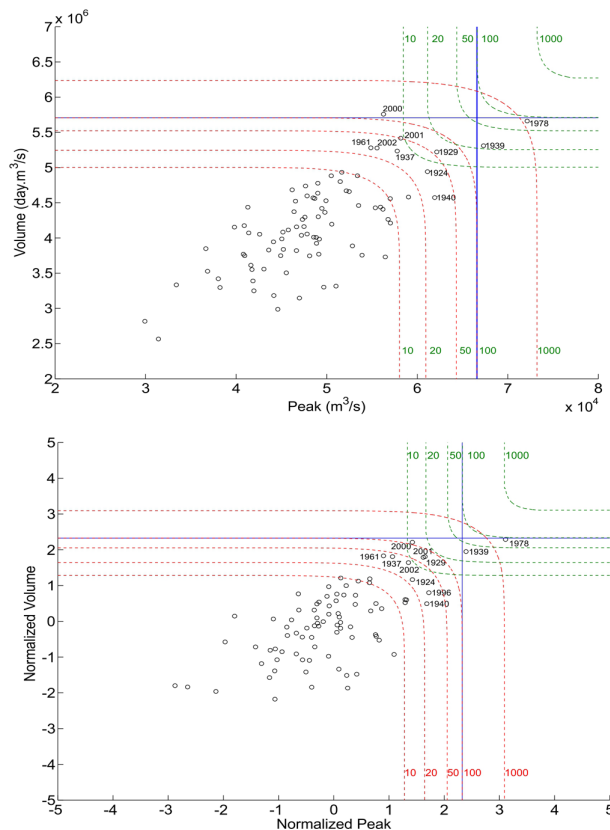
**Fig. 9.** Observed pairs of *Q* and *V* and events of different return periods (10, 20, 50, 100, 1000) as defined by the stationary (top) and non-stationary (bottom) copula for the "AND" (red) and "OR" (green) definition. The black circles indicate the observed data, with the largest events indicated by the year of occurrence. Note that for the non-stationary analysis normalized data are presented in order to eliminate the time shifting effect, thus enabling the plotting of the observed data in a single plot.
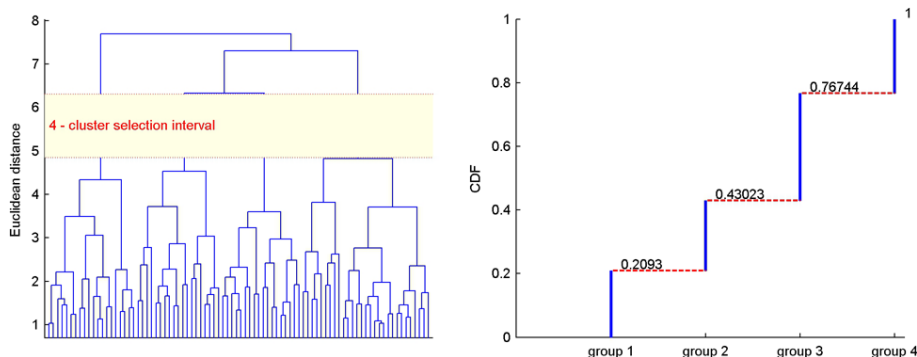
**Fig. 10.** (left) Dendrogram illustrating the cluster analysis for grouping hydrographs using the ward algorithm; (right) and cumulative frequency of the 4 chosen clusters.
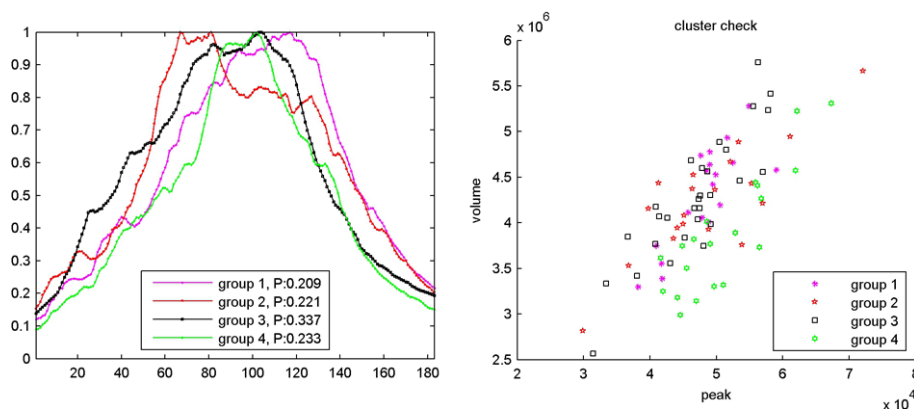
**Fig. 11.** (left) mean normalized hydrographs of the four-cluster solution; (right) distribution of the four clusters in the scatterplot of flood peak and volume.
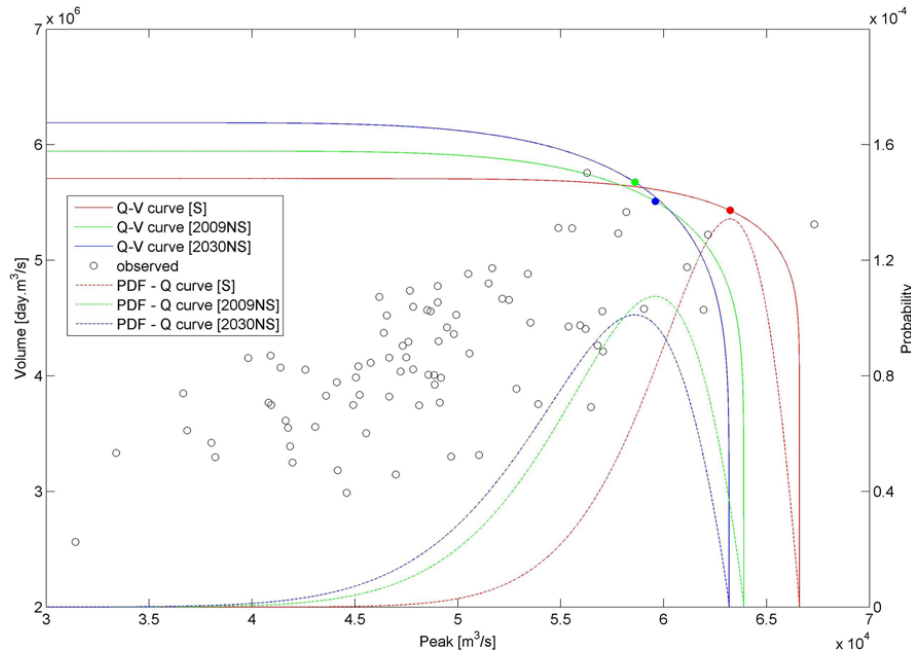
**Fig. 12.** 100-yr return period curves (solid), corresponding pdf's of the discharge along the curve (dashed) and the vertexes of the curve (point): red = stationary, green = non-stationary 2009, blue = non-stationary 2030.
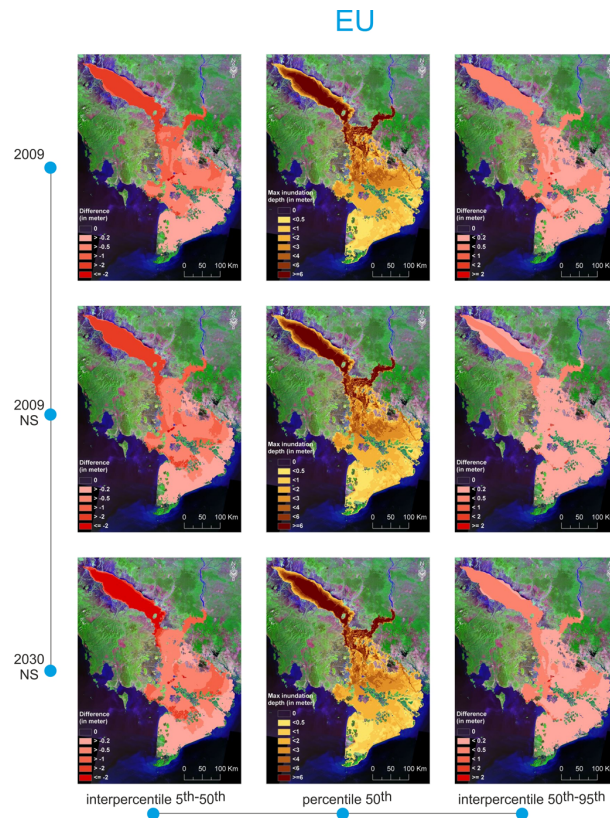
**Fig. 13.** Mekong Delta flood hazard maps for maximum inundation depth derived with the roughness parameter set "best EU" for the hydraulic model using different scenarios for the input: (top) 2009, (middle) 2009 – non-stationary approach, (bottom) 2030 – non-stationary approach.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Fig. 14.** Mekong Delta flood hazard maps for maximum inundation depth derived with the roughness parameter set "best $F_2$" for the hydraulic model using different scenarios for the input: (top) 2009, (middle) 2009 – non-stationary approach, (bottom) 2030 – non-stationary approach.
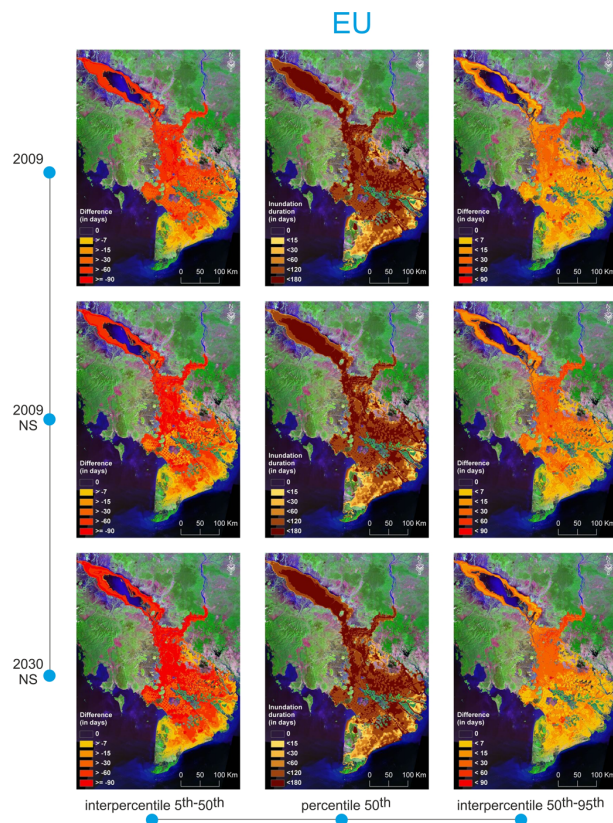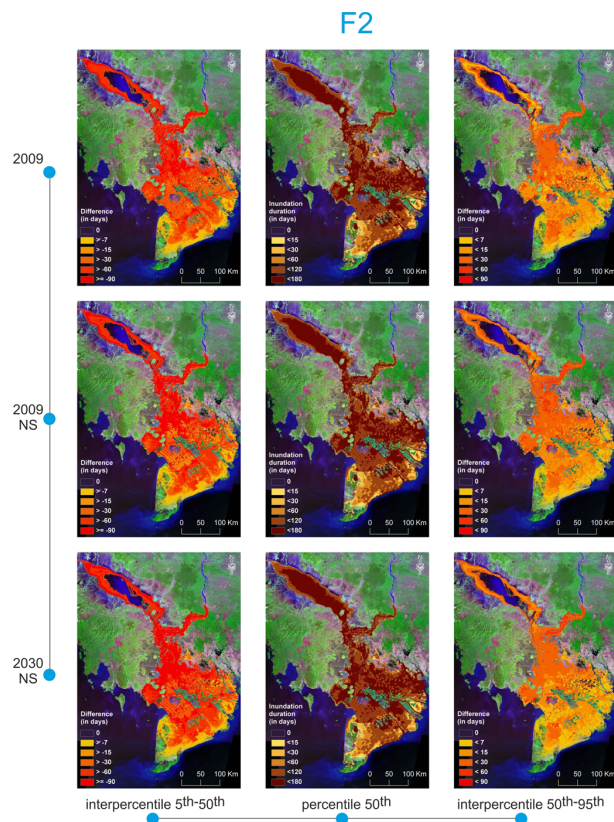
**Fig. 15.** Mekong Delta flood hazard maps for inundation duration derived with the roughness parameter set "best EU" for the hydraulic model using different scenarios for the input: (top) 2009, (middle) 2009 – non-stationary approach, (bottom) 2030 – non-stationary approach.

**Flood hazard in the Mekong Delta**

N. V. Dung et al.



**Fig. 16.** Mekong Delta flood hazard maps for inundation duration derived with the roughness parameter set "best $F_2$" for the hydraulic model using different scenarios for the input: (top) 2009, (middle) 2009 – non-stationary approach, (bottom) 2030 – non-stationary approach.