Real-time forecasting of typhoon inundation extent in a partially-gauged area through the integration of ARX-based models and a geographic information system

Huei-Tau Ouyang [1], Yi-Chun Chen[2]

[1]Department of Civil Engineering, National Ilan University, Yilan city, Yilan County 260, Taiwan

[2]First River Management Office, Water Resources Agency, Yilan City, Yilan County 260, Taiwan

Correspondence to: Huei-Tau Ouyang (htouyang@niu.edu.tw)

Abstract. This study presents a methodology for forecasting the extent of inundation and depth of distribution during typhoons in real-time. The proposed approach involves the construction of ARX and ARMAX models capable of predicting water-levels at the locations of on-site gauging stations and representative points located at the outlets of the sub-areas obtained by terrain analysis using a geographic information system. The models are constructed based on historical typhoon data and the results of numerical simulations related to inundation. A database comprising layers of inundation maps related to water-levels in each sub-area based on the assumption of flat-water and the digital elevation model (DEM) of the area were assembled prior to the typhoon. Water-levels during the typhoon are forecast using the constructed models, whereupon inundation sub-maps associated with the forecasted water-levels are extracted from the database. The resulting inundation map is comparable to that obtained using Synthetic Aperture Radar. Processing can be conducted in real-time and requires very little computational resources. This provides valuable lead time in which to conduct efforts aimed at damage mitigation during a typhoon.

1 Introduction

Every summer, Taiwan faces the threat of typhoons with heavy rainfall, which can result in serious inundation disasters in low-lying areas. Severe inundation can cause property loss and threaten human life. Mitigating damage from typhoons requires a rapid and reliable means to forecast the extent of inundation and water depth distribution during the typhoon. Extensive research has been conducted on inundation forecasting. Romanowicz and Beven (1998) developed a methodology for predicting space and time varying probabilities of inundation on a flood plain, based on Monte Carlo simulations of a distributed quasi-two-dimensional flood routing model using known levels at sites along the reach. Thirumalaiah and Deo (1998) developed a real-time flood forecasting model based on a neural network trained using a selected sequence of previous flood observations at a specific location. Toth et al. (2000) compared the advantages and limitations of ARMA, ANN, and the non-parametric nearest-neighbours method in forecasting rainfall runoff, concluding that time-series analysis is far more accurate than simple rainfall prediction of a heuristic nature. Nayak et al. (2005) proposed a real-time flood forecasting model based on fuzzy computation and concluded that the recursive use of a one-step-ahead forecast model to predict flow using longer lead times produces results better than those achieved using

1

independent fuzzy models to forecast flows at various lead times. Neal et al. (2007) proposed an approach to the forecasting of flood inundation based on a one-dimensional hydraulic flow simulation model updated using real-time data within an ensemble Kalman filtering framework. Romanowicz et al. (2008) presented a data-based mechanistic methodology to derive the non-linear dependence between water

5 levels measured at gauging stations along a river. Tang et al. (2009) developed a system for forecasting tsunami inundation based on direct observations of tsunamis and numerical modelling. Leedal et al. (2010) proposed a method for the real-time communication of forecasted flood levels and inundation-related information through the incorporation of model uncertainty analysis, data assimilation algorithms, two-dimensional inundation modelling, and data visualization. Pan et al. (2011) proposed a rainfall-

10 inundation forecasting model that works in real-time using a hybrid neural network based on a synthetic database of inundation potential.

As shown in the above literature review, one practical means of gathering on-site inundation information involves the construction of a surveillance network comprising gauging stations; however, financial constraints commonly limit the number of gauging stations, which restricts the range of surveillance that

15 can be maintained by the network. Furthermore, surveillance networks are able to model the inundation situation only as it occurs, without the ability to forecast the future events crucial to the process of mitigating damage. Numerical models based on physical laws have shown promise with regard to prediction; however, this type of model involves heavy time-consuming computation, which precludes the provision of information in real time with sufficient lead time to allow for the timely implementation

20 of countermeasures during a typhoon. Block-box modelling is based on the analysis of characteristics between input and output data to produce models capable of representing the relationships among inputs and outputs. Models such as Auto-Regressive with Exogenous Inputs (ARX) and Auto-Regressive Moving Average with Exogenous Inputs (ARMAX) are examples of this type of approach. These methods provide advantages with regard to computational speed and the results are generally comparable

25 to physical models; however, the fact that this approach relies heavily on observational data means that it is inapplicable in ungauged areas.

In the present study, we proposed a methodology integrating the advantages of each of these three approaches with the aim of achieving real-time forecasting as to the extent of inundation and depth distribution in partially gauged areas during typhoons. The study extended Ouyang (2016) by the

30 following improvements: (1) A geographic information system (HEC-GeoHMS) is incorporated in the proposed methodology to analyse the digital elevation model (DEM) of the study area and to divide the area into sub-areas; (2) The methodology proposed in the present study is capable of not only forecasting future inundation-levels but also the extent of inundation in the area by incorporating DEM and the flat-water assumption; (3) The rainfall water-level models in the present study are improved by taking in the

35 simulated inundation-level data from the inundation potential database (IPDB) as inputs, and thus are applicable for locations where no observation data is available; and (4) In the present study, not only the gauged locations but also the ungauged representative points (outlets of the sub-areas) are modelled. This paper is arranged as follows. A description of the proposed methodology is presented in Section 2. An example implementation of the proposed method is presented in Section 3. Section 4 outlines the

40 performance of the proposed rainfall and water-level relationship models. We also compare the resulting

inundation map with one obtained using Synthetic Aperture Radar. In the final section, conclusions are drawn based on the results.

2 Methodology

The proposed method proceeds through two stages: the preparation stage and the typhoon attack stage.

5   In the preparation stage, three types of data are collected: (1) historical rainfall and water-level records from previous typhoon events obtained at on-site gauging stations, (2) the results of numerical inundation simulations, and (3) a digital elevation model (DEM) of the area. These data are used to equip three major components for inundation forecasting: (1) Construction of a rainfall and water-level relationship model for each of the gauging sites; (2) Division into sub-areas and construction of a water-level model for each

10   of the sub-areas; and (3) Construction of a database comprising inundation sub-maps for each sub-area with various inundation-related water-levels. In the typhoon attack stage, water-levels in the sub-areas are forecast using the models constructed based on the on-site information transmitted from gauging stations, whereupon inundation sub-maps are extracted from the database for each of the sub-areas in accordance with the forecasting of inundation water-levels. Overlaying the inundation sub-maps makes

15   it possible to predict the extent of inundation and water depth distribution. Figure 1 presents a flow chart of the proposed methodology. Each part in the procedure is detailed in the following.

2.1 Relationship between rainfall and water-level at gauging sites

Due to meteorological and geographical conditions in Taiwan, inundations in low-lying areas usually occur soon after typhoon attacks. Physical inundation models are seldom fast enough to provide adequate

20   lead time in indicating areas requiring evacuation due to high risk of flooding. Considering the rapid responses required for disaster mitigation, it is crucial that forecasts on the extent of inundation and depth distribution be obtained as soon as possible. This study adopted black-box models, which have proven particularly advantageous in this regard for their ability to provide reasonable predictions in a timely manner (Karlsson and Yakowitz, 1987).

25   Block-box modelling to simulate the rainfall-runoff relationship has been investigated extensively (for example, Shamseldin and O'connor, 1999; Gautam and Holz, 2001; Talei et al., 2010). These models are fast and the results are often comparable with physical models. This study employed ARX and ARMAX models due to their simplicity and ease of integration with Geographic Information System (GIS). It is worth noting that in estimating the extent of inundation and water depth distribution, the water-level is a

30   greater concern than runoff. Thus, in our study the models are constructed to establish the relationship between rainfall and the water-level rather than runoff. It should also be noted that real-time forecasting during a typhoon attack prevents the normalization of data on rainfall and water-level according to the maximum and minimum prior to regression. This has commonly been done in the literature; however, such information is unavailable during the typhoon event. Instead, the proposed models take in raw data

35   related to rainfall and water-levels and output the predicted water-levels for the next time-step. This means that the calibrated models are applicable in real time during the typhoon attack.

Hydrology and
Earth System
Sciences
Discussions

The ARX model was adapted from the Auto-Regressive Model (AR) proposed by Yule (1927), which is rooted in time series analysis using past data to predict future values. The AR model was extended to the ARX model through the adoption of exogenous variables. The ARX model has the following form:

$$H(t) + a_1 H(t-1) + \cdots + a_{n_a} H(t-n_a) = b_1 R(t-n_k) + \cdots$$

5
$$+ b_{n_b} R(t - n_b - n_k + 1) + e(t) \tag{1}$$

where H is the time series variable; R denotes the exogenous time series input; n_a and n_b are numbers indicating previous H and R respectively incorporated in the model; n_k denotes the response delay of R affecting y; a_1 through a_{n_a} and b_1 through b_{n_b} are adjustable parameters for H and R, respectively, which can be calibrated using regression. In this study, system output H is the water-level at each gauging

10 site, and exogenous input R is the increment of rainfall in each time step. The form of the model (or "model structure") is designated by the numbers n_a, n_b, and n_k. It should be noted that n_k is set at zero in this study to enable the inclusion of previous rainfall data $R(t-1)$ through $R(t-n_b)$ as well as rainfall forecasting data $R(t)$ provided by Central Weather bureau (CWB) of Taiwan.

The ARMAX model is another type of linear black-box model, which is an extension of the Auto-

15 Regressive Moving Average (ARMA) model. ARMA model was first proposed by Box and Jenkins (1976) as a combination of the AR model and the MA (Moving Average) model proposed by Slutsky (1937). ARMA model was then extended as the ARMAX model through the incorporation of exogenous inputs in the following form:

$$H(t) + a_1 H(t-1) + \cdots + a_{n_a} H(t-n_a) = b_1 R(t-n_k) + \cdots$$

20
$$+ b_{n_b} R(t - n_b - n_k + 1) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c) + e(t) \tag{2}$$

where e is white noise; n_c is the number of previous e incorporated into the model; and c_1 through c_{n_c} are adjustable parameters for e.

To determine the structure of the model for each of the gauging sites, preliminary studies are used to guide the development of a number of candidate models of various structures. The performance of each

25 model is then evaluated using the various indices described in Section 2.3, whereupon the model best suited for each of the gauging sites is selected according to the evaluation results.

2.2 Rainfall and water-level relationship model for sub-areas

Due to economic constraints, the number of on-site gauging stations is often limited and the area covered by each gauging station is excessively wide, such that measured inundation data reflect only a small

30 portion of a given area close to the instrument. To expand the range of gauging data, a Digital Elevation Model of the area is collected and analyzed using the Geospatial Hydrological Modeling Extension (HEC-GeoHMS, Doan, 2000). HEC-GeoHMS was developed by the Hydrologic Engineering Center in US Army Corps of Engineers in collaboration with ESRI ArcGIS (Maidment, 2002). It is capable of delineating sub-areas in an area by analyzing the direction and accumulation of flow in cell-wise manner.

35 For each sub-area, a representative point located at the outlet of the area is identified using features in ArcGIS. The ARX and ARMAX techniques are then used to construct models for the relationship

between rainfall and the water-level at these points. However, these models differ from previous models in two details: (1) A lack of measurement data pertaining to the water-level at these locations necessitates the use of computational data from numerical inundation simulation for the construction of models; (2) Computational data pertaining to the water-level at gauging sites are also incorporated into the models as exogenous inputs. This means that the resulting models reference rainfall forecasting data as well as measurement data pertaining to the water-level obtained on-site at a later stage during the typhoon attack. In other words, the water-levels predicted for these representative points incorporate all of the information from the numerical simulations as well as the forecasted rainfall data and observations of water-levels obtained on-site.

For the numerical simulation of inundation, we employed the inundation potential database (IPDB), which was constructed by the Taiwan Water Resources Agency (WRA) in 2007. The IPDB comprises simulation data pertaining to inundation to assist in the division of watersheds under 22 rainfall scenarios: daily rainfall of 200mm, 350mm, 450mm, 600mm; two-day rainfall of 450mm, 600mm, 750mm, 900mm; three-day rainfall of 750mm, 900mm, 1050mm, 1200mm; and frequent rainfall events with return periods of 1.1, 2, 5, 10, 20, 25, 50, 100, 200, 500 years. Numerical simulation involves dividing each watershed into an upstream mountainous catchment and downstream alluvial plain. The HEC-HMS model (Hydrological Modelling System, Feldman, 2000) was used to calculate the discharge flowing down from the upstream catchments. For the downstream areas, two-dimensional overland-flow routings (Hsu et al., 2002) were utilized for surface inundation simulations incorporated with one-dimensional dynamic channel-flow routings (Hsu, 1992) for main channel flow computations and the SWMM model (Huber et al., 1988) for sewer flow routings. In the present study, the inundation data obtained from the simulation of quantitative rainfalls in IPDB was used to construct and calibrate models of the relationship between rainfall and water-level at the representative points, and inundation data related to rainfall frequency was used for calibration of the model.

The models used for representative points were incorporated with data related to the water-level at gauging sites, which made it possible to assimilate information obtained on-site into the forecasting of water-level at these locations during a typhoon attack. To identify appropriate gauging sites for incorporation at each of the representative points in the model, we analysed correlations between water-level hydrographs at the representative points and gauging sites. For each representative point, we selected the corresponding gauging with the highest correlation. We then incorporated water-level data from the corresponding gauging site within the model as an exogenous input, as shown in the following equations:

$$H(t) + a_1 H(t-1) + \cdots + a_{n_a} H(t-n_a) = b1_1 R(t-n_{k1}) + \cdots + b1_{n_{b1}} R(t-n_{b1}-n_{k1}+1)$$

$$+ b2_1 H_2(t-n_{k2}) + \cdots + b2_{n_{b2}} H_2(t-n_{b2}-n_{k2}+1) + e(t) \qquad (3)$$

$$H(t) + a_1 H(t-1) + \cdots + a_{n_a} H(t-n_a) = b1_1 R(t-n_{k1}) + \cdots + b1_{n_{b1}} R(t-n_{b1}-n_{k1}+1) +$$
$$b2_1 H_2(t-n_{k2}) + \cdots + b2_{n_{b2}} H_2(t-n_{b2}-n_{k2}+1) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c) + e(t) \qquad (4)$$

where Eqs. (3) and (4) respectively represent the ARX and ARMAX models; H denotes water-level at the representative point; R is the rainfall increment and H_2 is data related to water-level from the

5

associated gauging station; n_{b1} and n_{b2} are the numbers of past data included in the model associated with R and H_2, respectively; and n_{k1} and n_{k2} represent the response delay of R and H_2 to H; and $b1_1$ through $b1_{n_{b1}}$ and $b2_1$ through $b2_{n_{b2}}$ are adjustable parameters associated with R and H_2, respectively.

2.3 Evaluation of the models

5 We proposed a number of candidate models for each gauging site and representative point, and the performance of each candidate model is evaluated using the following four indices.

(1) Efficiency Coefficient (CE)

CE was proposed by Nash and Sutcliffe (1970) to assess the predictive ability of a hydrological model. CE is defined as follows:

10
$$CE = 1 - \frac{\sum_{t=1}^{n}[H_{obs}(t) - H_{est}(t)]^2}{\sum_{t=1}^{n}[H_{obs}(t) - \bar{H}_{obs}]^2} \tag{5}$$

where H_{obs} and H_{est} represent the observed and estimated water-levels, respectively; $\bar{H}_{obs}$ is the averaged observation of water-level; n is the number of data points. CE quantifies the goodness of fit between the observed and estimated hydrographs of the water-levels. Value of CE closer to 1 indicates a model with better fit.

15 (2) Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n_t}(H_{obs}(t) - H_{est}(t))^2}{n}} \tag{6}$$

RMSE represents the mean error between the observed and estimated hydrographs. A smaller RMSE value indicates a better model.

(3) Error in the stage of peak water-level (ESP)

20
$$ESP = \frac{H_{p,est} - H_{p,obs}}{H_{p,obs}} \times 100\% \tag{7}$$

where $H_{p,obs}$ and $H_{p,est}$ denote the observed and estimated peaks in the water-level, respectively. ESP is used to quantify the error between the observed peak in the water-level and the model estimation. A positive ESP indicates an over estimation by the model, while a negative ESP indicates under estimation. A smaller absolute ESP value indicates a better fit.

25 (4) Error in time of peak arrival (ETP)

$$ETP = T_{p,est} - T_{p,obs} \tag{8}$$

where $T_{p,est}$ and $T_{p,obs}$ respectively denote the predicted and observed arrival times of the peaks in water-levels. A positive ETP indicates that the predicted arrival time is early. From a warning perspective, a positive ETP would be preferable to a negative ETP, due to the fact that it would provide additional time

30 for the implementation of actions intended for the mitigation of damage.

2.4 Construction of database comprising inundation sub-maps

We employed the flat-water model (Zerger, 2002, Zerger et al., 2002, Chen et al., 2009), which assumes a planar water surface over the inundated area, for each sub-area in the construction of inundation sub-maps. The flat-water model has proven suitable for small flat environments (Zerger et al., 2002),

5 particularly in cases of typhoon inundation where the flow velocity is almost at rest. Inundation sub-maps are constructed by overlapping the DEM of the area with various pre-defined flat-water surfaces, wherein the elevation is increased in increments of 1cm starting from the bottom of the area. These pre-determined inundation sub-maps then serve as a database for later use in typhoon attacks.

In the typhoon attack stage, we first forecast the water-levels at gauging sites using the models based on

10 rainfall forecast data from CWB and measurements of water-level transmitted from on-site gauging stations. The water-level at each of the representative points is then predicted by the model based on rainfall data and the water-level data from the associated gauging site. One-step-ahead forecasts are processed recursively to gain higher lead time. Inundation sub-maps associated with the sub-areas are extracted from the database according to the water-levels predicted for each of the representative points

15 for the assembly of a complete inundation map covering the entire area. In extracting the sub-maps, the predicted water-level is rounded off to the closest water-level found in the database with precision of 1 cm. A complete inundation map of the area is then assembled by layering all of the inundation sub-maps. The process is repeated regularly (updated) throughout the typhoon attack in accordance with the transmission of water-level data from on-site gauging stations. Processing can be conducted in real time

20 due to the fact that all of the models are calibrated and the inundation sub-maps are pre-determined in the preparation stage. The computation required for the models and the extraction of inundation sub-maps from the database require almost no time at all, thereby providing additional lead time with which to implement actions for the mitigation of damage during the event.

3 Practical application

25 The methodology developed in this study was implemented as a practical application to obtain early warnings of typhoon inundation in an inundation-prone area in Taiwan. The Xinnan area studied by Ouyang (2016) was selected as the study area. The area covers 6.5 km2 and the elevation of terrain in the area averages approximately 2 meters only. Due to the low elevation, the area is notoriously prone to inundation under typhoon attack, which cause a serious loss of property and threaten the life of local

30 residents. In 2011, the Surveillance Network for Typhoon Inundation in the Xinnan Area (SNTIX) was constructed in the area to provide on-site inundation information and forecast potential inundation situations during typhoon attacks. The SNTIX comprises four water-level gauging stations (Zhongnanxing, Xinnan, Sijie, and Meifu) with radio transmission in conjunction with a data transmission interface to receive rainfall observation data and forecasts from the QPESUMS

35 (Quantitative Precipitation Estimation and Segregation Using Multiple Sensor, Gourley et al., 2002) of the Central Weather Bureau (CWB) in Taiwan. Detailed descriptions about SNTIX and the Xinnan area can be found in Ouyang (2016).

Hydrology and
Earth System
Sciences

Open Access

EGU

Discussions

Obtaining early warnings requires that SNTIX be able to forecast future variations in inundation situations in the target area over the following few hours in order to identify locations for evacuation and implement actions for the mitigation of damage in-advance in order to prevent catastrophes. We integrated the proposed methodology with SNTIX to facilitate inundation forecasting. The results are outlined in the following.

4 Results and discussion

4.1 Modelling of rainfall water-level at gauging stations

We employed data of typhoon events collected by SNTIX in the construction of models for the representation of the relationship between rainfall and water-level at the gauging sites. A preliminary study based on ARX and ARMAX indicated a number of candidate models for the gauging sites in the Xinnan area to simulate the relationship between rainfall and water-level. Table 1 lists the candidate models for each of the gauging sites, in which arx and amx denote ARX and ARMAX models, respectively, and the number signifies the structure of the model. The first figure denotes the number of previous H incorporated in the model, the second figure designates the number of past R, and the third figure represents the time lag of R. For example, arx440 indicates that the model is based on ARX, comprising 4 terms in H starting from $H(t-1)$ to $H(t-4)$; and 4 terms in R with 0 lag starting from $R(t)$ to $R(t-3)$. Similarly, amx4720 indicates an ARMAX model with 4 terms in y, 7 terms in R with 0 lag, and 2 terms in e. As mentioned previously, in all of the candidate models, the lag of R is specified as 0 in order to incorporate forecasted rainfall data $R(t)$ from QPESUMS in the models.

The candidate rainfall water-level models listed in Table 1 were calibrated using typhoon event data recorded by SNTIX. The events with the highest, the lowest, and the median cumulative rainfalls were selected for model calibration, and the other events in the database were used for model validation. The performance of each model was evaluated using the indices of CE, RMSE, ESP, and ETP in which we compared model predictions with three-hour lead times against actual measurement data.

Table 2 shows the event-averaged model performance for calibrations. As seen in the table, the calibrated CE of each model exceeds 0.8 with RMSE values of 0.03 to 0.06 meters, indicating that the models performed well when compared with measurement data. The predictive ability of the models in estimating the peak water- stage and associated arrival time are indicated by ESP and ETP, respectively. The ESPs for the calibrated models are within a range of approximately -2.0% to 0.8%, and the ETPs are -1.5 to 1.2 hours, both of which are satisfactory. The ETPs of some of the models are somewhat high, which indicates that the times predicted for the arrival of peak water-levels was somewhat early. Nonetheless, this may be advantageous from the perspective of early warning, in that it provides additional time in which to implement actions for the mitigation of damage.

As shown in Table 2, model arx170 for the Zhongnanxing station appears to have the highest CE and the lowest RMSE. It also has the lowest ESP in the prediction of peak water-levels. It produced an ETP of 1.2 hours; however, this is not considered a drawback in terms of early warning. In the case of the Xinnan station, model amx8380 was selected for use in SNTIX, due to the fact that it has the highest CE and lowest RMSE. For the Sijie station, model amx7790 presented the best scores with regard to CE, RMSE,

and ESP and was therefore selected as the optimal model for the station. For the Meifu station, model amx1450 presented the highest score on CE and was selected for use in SNTIX.

Figure 2(a) through (d) present the calibrated water-level predictions with 3hr lead times as well as data measured during typhoon Saola at four gauging sites. As shown, most of the predictions are in good

5   agreement with the measurements. It is worth noting that the selected models also provide good predictions with regard to peak water-levels, which are considered the most important information in terms of mitigating damage. Predictions by the model for Meifu station related to the arrival of the peak water-level were slightly late. This may be due to the fact that this station is located in the most downstream area in which variations in the water-level may be affected by tidal effects that are not

10   accounted for in the model.

The selected models were then validated using the rest of the typhoon events in the database. Table 3 presents the event-averaged performance of the model with 3hr lead times. The models for the Xinnan, Sijie, and Meifu stations achieved CE scores exceeding 0.8, thereby validating the good overall prediction results during typhoon events. As shown in the table, the ESP scores for the prediction of peak

15   water-levels are generally within -0.6% to 2.3%. The negative ETPs indicate that the predictions of peak arrival time are late; however, they are still less than 3 hours, which is considered sufficient for the provision of early warnings.

Figure 3(a) through (d) present a comparison of the validated water-level predictions with 3hr lead times against data recorded during typhoon Usagi, showing good agreement between the predictions and

20   measured data. The models exhibited satisfactory ability in predicting peaks and the time of peak arrival. The predictions for the peak arrival times at Xinnan and Meifu stations were slightly late; however, the deviations are shorter than the lead time used in forecasting.

4.2 Inundation sub-maps and rainfall water-level modelling for sub-areas

A high-resolution digital elevation model (DEM) of the study area was produced using airborne LIDAR

25   with precision of 0.5×0.5 m2. Due to constraints in project funding, we were limited to four on-site gauging stations, each covering an area of approximately 1~2 km2. To increase the accuracy of inundation forecasting, we analyzed the DEM of the study area using HEC-GeoHMS and divided the area into 22 sub-areas, as shown in Figure 4. Representative points located at the outlets of the sub-areas were also identified, as shown in the figure. It should be noted that only nine representative points were

30   used because some of the sub-areas share the same outlets and four gauging sites are also located at the outlets of sub-areas. By dividing the study area into smaller portions, the area covered by each sub-area was greatly reduced to only 0.2~0.3 KM2, thereby fulfilling the assumption of flat-water. Inundation sub-maps were constructed by overlapping the DEM of each sub-area with various flat-water planes produced with 1cm increments in surface elevation starting from the lowest elevation of the sub-area.

35   These pre-determined inundation sub-maps were stored in a GIS database for further use.

We also characterized the relationship between rainfall and water-level at the nine representative points using black-box models of ARX and ARMAX. These representative points are ungauged due to a lack of available measured data; therefore, we used data from IPDB numerical simulations for the construction of models. Water-stage hydrographs for the nine representative points and the four gauging sites were

extracted from IPDB to serve as a database for the construction of a model used to characterize the relationship between rainfall and the water-level, in which the inundation data related to the quantitative rainfall is used for calibration of the model and frequency inundation data is employed for model validation.

5   To identify the gauging site associated with each of the representative points, we conducted correlation analysis on the computed water-levels between the representative points and the gauging sites. Table 4 presents the correlation coefficients averaged from all of the events from IPDB, in which those with the highest scores are marked using an asterisk. Having identified the associated gauging sites, a number of model structures are proposed for each of the representative points using Eqs. (3) and (4), in which

10   incremental rainfall data is incorporated with data related to the water-level from the associated gauging sites as exogenous inputs. The performance of each model was evaluated using the four indices (CE, RMSE, ETP, and ESP) with lead times of 3 hours, and the ones with the highest average CE in model calibration are presented in Table 5. The structure of the model is designated by the name of the model. For example, arx4[5 3][1 0] refers to an ARX model incorporating four data points of y, five data points

15   from the associated gauging station, and three terms from the rainfall data, with lag time of 1 for the data from the associated gauging station and lag time of 0 from the rainfall data. In the case of the ARMAX model, the figure between the two brackets denotes the number of white noise terms incorporated in the model. As shown in the table, the average CE values for the selected models are approximately 0.8~0.9 for calibration and 0.75~0.85 for validation. The average ESP values for model calibration and validation

20   are within approximately $\pm5\%$, and the ETPs are approximately ±2 hours. These are satisfactory prediction results for water-levels and peaks. Figure 5 presents a comparison between model predictions and IPDB data at representative point P8. As shown in Figure 5 (a), there is good agreement between the prediction results and measurement data for model calibration. For model validation, little late prediction is observed as seen in Figure 5 (b).

25   **4.3 Inundation forecasting during typhoon attacks**

The construction of models for gauging sites and representative points enables SNTIX to predict variations in water-levels at specific locations for the following few hours during a typhoon attack. These predictions are based on rainfall forecasting data from QPESUMS and water-level data transmitted from gauging stations. The pre-determined inundation sub-maps associated with the predicted water-levels at

30   representative points were obtained from the database for the assembly of inundation map in that area. Figure 6 (a) illustrates the forecasted extent of inundation and water depth distribution with a lead time of 3-hr for typhoon Saola, which occurred at PM 17:50 on August 2nd 2012. For comparison, the extent inundation as determined by Synthetic Aperture Radar from TerraSAR-X (Stang et al., 2006) is presented in Figure 6 (b). The extent of inundation predicted by the proposed methodology (4.33 km2) is similar

35   to that of TerraSAR-X (4.02 km2), which represents deviation of approximately 8%. This deviation may be attributed to the assumption of flat-water, which departs somewhat from the actual distribution of water surfaces in the area. It should be noted that the proposed forecast was obtained in real time during the event, while the result from TerraSAR-X was acquired after the event. Furthermore, depending on the weather and cloud conditions during typhoon events, inundation data is not necessarily available from

TerraSAR-X. To enable the rapid identification of potential inundation regions in the area, the proposed methodology is able to provide essential information with ample lead time for the implementation of actions for the mitigation of damage during typhoon attacks.

5 Conclusion

5   A simple and practical methodology for the real-time forecasting of inundation extent and water depth distribution in flat areas resulting from typhoon attacks has been developed. The approach involves the division of the target area into sub-areas and the incorporation of rainfall data, water-level observations, and results of numerical inundation simulations to construct ARX/ARMAX models representing the relationship between rainfall and water-level at gauging sites and ungauged representative points located
10   at the outlets of the sub-areas. Inundation sub-maps are pre-constructed for each sub-area based on the assumption of flat-water with various pre-defined inundation water surface elevations. During typhoon attacks, the water-levels at gauging sites and representative points are predicted by the models and an inundation map is assembled by extracting the sub-maps associated with the predicted water-levels from the database. The proposed methodology was implemented with a surveillance network in an inundation-
15   prone area in Taiwan to enable rapid forecasting and the identification of potential regions of inundation in the area during typhoon attacks. The construction of ARX/ARMAX models for gauging sites and representative points achieve a satisfactory degree of fit between events and peaks. Forecasts related to the extent of inundation show a promising degree of agreement with the data obtained from Synthetic Aperture Radar. The ability to forecast the situation of on-site inundation in real time renders the
20   proposed methodology a fast and reliable approach to decision-making for actions pertaining to the mitigation of damage during typhoons.

Acknowledgement

25   ### References

Box, G. E., & Jenkins, G. M. (1976). Time series analysis: forecasting and control, revised ed. Holden-Day.

Chen, J., Hill, A. A., & Urbano, L. D. (2009). A GIS-based model for urban flood inundation. Journal of Hydrology, 373(1), 184-192.

30   Doan, J. (2000). Geospatial hydrologic modeling extension HEC-GeoHMS user's manual version 1.0. Davis, California.

Feldman, A. D. (2000). Hydrologic modeling system HEC-HMS: technical reference manual. US Army Corps of Engineers, Hydrologic Engineering Center.

Gautam, D., & Holz, K. (2001). Rainfall-runoff modelling using adaptive neuro-fuzzy systems. Journal
35   of Hydroinformatics, 3, 3-10.

11

Gourley, J. J., Maddox, R. A., Howard, K. W., & Burgess, D. W. (2002). An exploratory multisensor technique for quantitative estimation of stratiform rainfall. Journal of Hydrometeorology, 3(2), 166-180.

Hsu, M. H. (1992). Simulation of inundation with overflow on Levee along Keelung River. In The CCNAA-AIT Joint Seminar on Prediction and Damage Mitigation of Meteorologically Induced Natural Disasters.

Hsu, M. H., Chen, S. H., & Chang, T. J. (2002). Dynamic inundation simulation of storm water interaction between sewer system and overland flows. Journal of the Chinese Institute of Engineers, 25(2), 171-177.

Huber, W. C., Dickinson, R. E., Barnwell Jr, T. O., & Branch, A. (1988). Storm water management model, version 4. US Environmental Protection Agency, Environmental Research Laboratory.

Karlsson, M., & Yakowitz, S. (1987). Rainfall-runoff forecasting methods, old and new. Stochastic Hydrology and Hydraulics, 1(4), 303-318.

Leedal, D., Neal, J., Beven, K., Young, P., & Bates, P. (2010). Visualization approaches for communicating real-time flood forecasting level and inundation information. Journal of Flood Risk Management, 3(2), 140-150.

Maidment, D. R. (2002). Arc Hydro: GIS for water resources (Vol. 1). ESRI, Inc.

Nash, J., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology, 10(3), 282-290.

Nayak, P. C., Sudheer, K. P., & Ramasastri, K. S. (2005). Fuzzy computing based rainfall-runoff model for real time flood forecasting. Hydrological Processes, 19(4), 955-968.

Neal, J. C., Atkinson, P. M., & Hutton, C. W. (2007). Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. Journal of Hydrology, 336(3), 401-415.

Ouyang, H. T. (2016). Multi-objective optimization of typhoon inundation forecast models with cross-site structures for a water-level gauging network by integrating ARMAX with a genetic algorithm. Natural Hazards and Earth System Sciences, 16(8), 1897-1909.

Pan, T. Y., Lai, J. S., Chang, T. J., Chang, H. K., Chang, K. C., & Tan, Y. C. (2011). Hybrid neural networks in rainfall-inundation forecasting based on a synthetic potential inundation database. Natural Hazards and Earth System Science, 11(3), 771-787.

Romanowicz, R. J., Young, P. C., Beven, K. J., & Pappenberger, F. (2008). A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting. Advances in Water Resources, 31(8), 1048-1056.

Romanowicz, R., & Beven, K. (1998). Dynamic real-time prediction of flood inundation probabilities. Hydrological Sciences Journal, 43(2), 181-196.

Shamseldin, A. Y., & O'CONNOR, K. M. (1999). A real-time combination method for the outputs of different rainfall-runoff models. Hydrological Sciences Journal, 44(6), 895-912.

Slutzky, E. (1937). The summation of random causes as the source of cyclic processes. Econometrica: Journal of the Econometric Society, 105-146.

Stangl, M., Werninghaus, R., Schweizer, B., Fischer, C., Brandfass, M., Mittermayer, J., & Breit, H. (2006). TerraSAR-X technologies and first results. IEE Proceedings-Radar, Sonar and Navigation, 153(2), 86-95.

Talei, A., Chua, L. H. C., & Wong, T. S. (2010). Evaluation of rainfall and discharge inputs used by Adaptive Network-based Fuzzy Inference Systems (ANFIS) in rainfall–runoff modeling. Journal of Hydrology, 391(3), 248-262.

Tang, L., Titov, V. V., & Chamberlin, C. D. (2009). Development, testing, and applications of site-specific tsunami inundation models for real-time forecasting. Journal of Geophysical Research: Oceans (1978–2012), 114(C12).

Thirumalaiah, K., & Deo, M. C. (1998). Real-time flood forecasting using neural networks. Computer-Aided Civil and Infrastructure Engineering, 13(2), 101-111.

Toth, E., Brath, A., & Montanari, A. (2000). Comparison of short-term rainfall prediction models for real-time flood forecasting. Journal of Hydrology, 239(1), 132-147.

Yule, G. U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 267-298.

Zerger, A. (2002). Examining GIS decision utility for natural hazard risk modelling. Environmental Modelling & Software, 17(3), 287-294.

Zerger, A., Smith, D. I., Hunter, G. J., & Jones, S. D. (2002). Riding the storm: a comparison of uncertainty modelling techniques for storm surge risk management. Applied Geography, 22(3), 307-330.

Table 1 Structure of candidate model for the relationship between rainfall and water-level at gauging sites

Site	Structure of candidate model				
Zhongnanxing	arx110	arx120	arx130	arx140	arx170
Xinnan	amx3290	amx5790	amx8380	amx9290	amx9770
Sijie	arx440	amx4720	amx4890	amx4990	amx7790
Meifu	amx1240	amx1340	amx1350	amx1440	amx1450

Table 2 Event-averaged performance of calibrated candidate models for gauging sites (3-hr lead time)

Gauging station	Model	CE	RMSE (m)	ESP (%)	ETP (hr)
	arx110	0.906	0.050	0.81	0.8
	arx120	0.891	0.049	0.64	1.0
Zhongnanxing	arx130	0.872	0.047	0.50	1.0
	arx140	0.909	0.048	0.39	1.2
	arx170	0.912	0.045	0.36	1.2
	amx3290	0.815	0.039	0.05	0.2
	amx5790	0.865	0.067	-0.21	0.2
Xinnan	amx8380	0.895	0.037	0.29	0.2
	amx9290	0.825	0.041	0.29	0.2
	amx9770	0.877	0.039	0.27	-0.2
	arx440	0.905	0.049	0.72	0.8
	amx4790	0.845	0.057	-1.21	1.2
Sijie	amx4890	0.874	0.054	-1.24	0.2
	amx4990	0.892	0.052	-1.23	1.0
	amx7790	0.916	0.030	0.01	0.2
	amx1240	0.884	0.058	-0.37	-1.5
	amx1340	0.905	0.051	-1.93	-1.5
Meifu	amx1350	0.865	0.053	-1.65	-1.3
	amx1440	0.852	0.057	-2.01	-1.5
	amx1450	0.916	0.050	-0.50	-1.3

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

Table 3 Event-averaged model validation for gauging sites (3-hr lead time)

Gauging station	Model	CE	RMSE (m)	ESP (%)	ETP (hr)
Zhongnanxing	arx170	0.781	0.102	0.44	-1.2
Xinnan	amx8380	0.868	0.061	2.38	-1.8
Sijie	amx7790	0.833	0.071	-0.63	-0.2
Meifu	amx1450	0.847	0.060	-0.67	-2.5

Table 4 Averaged correlation coefficient between data related to water-level at representative points and gauging sites

Representative point	Gauging site			
	Zhongnanxing	Xinnan	Sijie	Meifu
P1	0.870	0.769	0.918*	0.857
P2	0.732	0.867*	0.863	0.703
P3	0.803	0.868*	0.758	0.801
P4	0.926	0.960*	0.819	0.810
P5	0.775	0.851*	0.786	0.687
P6	0.743	0.931*	0.877	0.683
P7	0.828*	0.767	0.784	0.793
P8	0.846*	0.803	0.815	0.793
P9	0.921*	0.885	0.866	0.878

Table 5 Event-averaged performance of the models for representative points (3-hr lead time)

Representative point	Selected model	Calibration				Validation			
		CE	RMSE (m)	ESP (%)	ETP (hr)	CE	RMSE (m)	ESP (%)	ETP (hr)
P1	arx4[5 3] [1 0]	0.860	0.040	0.05	0.2	0.818	0.066	0.08	-0.5
P2	arx3[3 5] [1 0]	0.901	0.072	-0.69	0.5	0.769	0.130	0.27	0.7
P3	arx5[2 6] [1 0]	0.907	0.021	0.22	-0.3	0.828	0.061	0.66	-0.8
P4	amx 4[6 4]3[3 0]	0.912	0.061	0.73	-0.5	0.793	0.085	4.45	-0.5
P5	arx5[4 3] [1 0]	0.907	0.068	0.20	0.2	0.888	0.069	0.99	0.8
P6	amx 5[5 4]3[1 0]	0.901	0.117	-0.14	-0.5	0.899	0.134	1.12	-1.2
P7	amx 3[6 3]2[2 0]	0.912	0.033	0.31	0.2	0.850	0.166	1.59	-0.5
P8	arx5[6 8] [3 0]	0.863	0.133	2.31	-0.3	0.796	0.142	2.91	0.5
P9	arx4[3 5] [1 0]	0.907	0.086	-0.14	-0.5	0.864	0.097	-1.70	-0.8

16
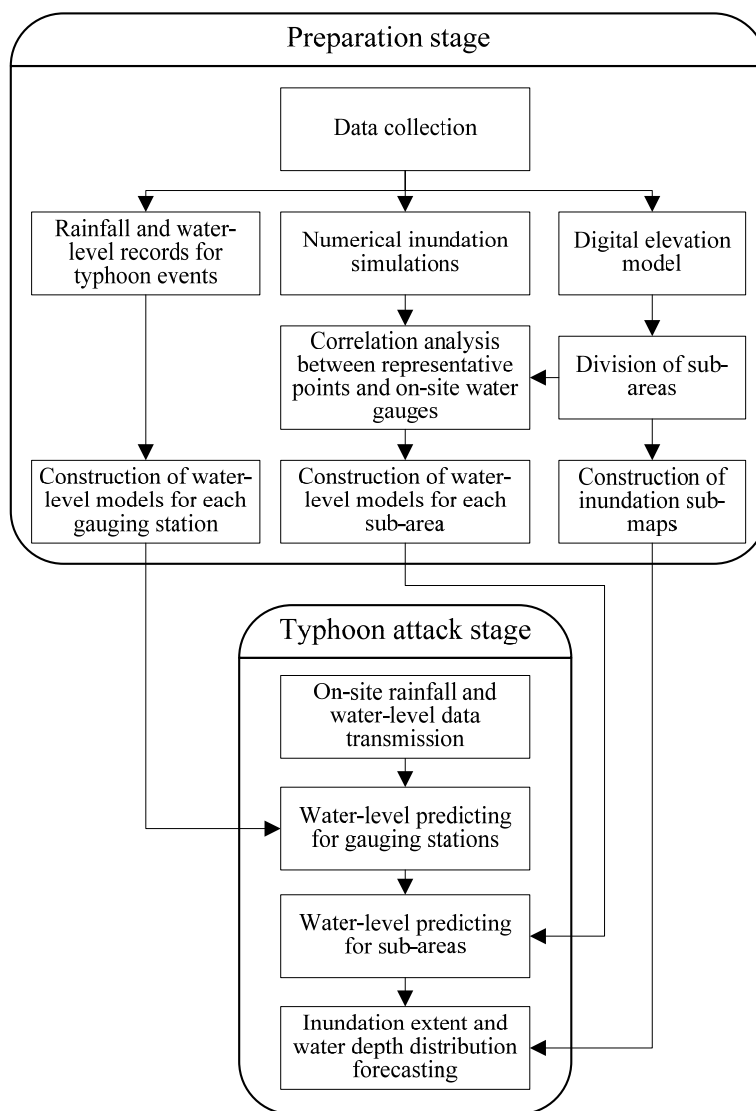
Figure 1 Flow chart illustrating proposed methodology
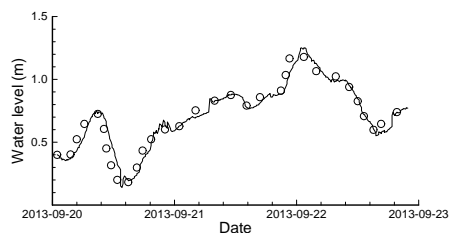
17
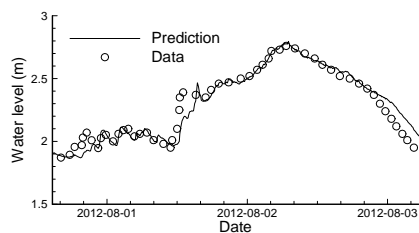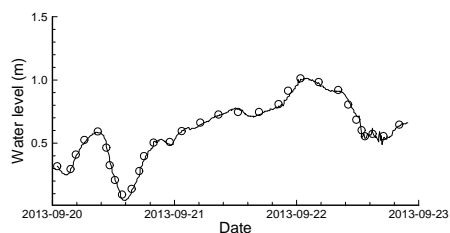
(a)
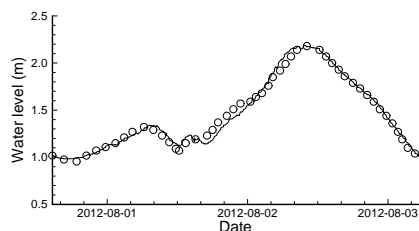
(b)
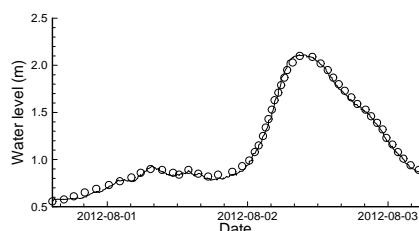
(c)
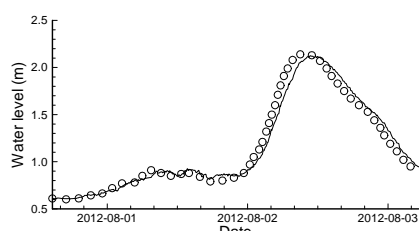
(d)

Figure 2 Comparison of calibrated water-level prediction (3-hr lead time) and measured data recorded during typhoon Saola: (a) Zhongnanxing station (b) Xinnan station (c) Sijie station (d) Meifu station
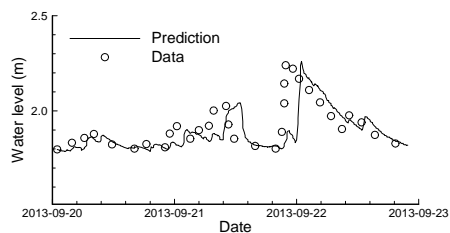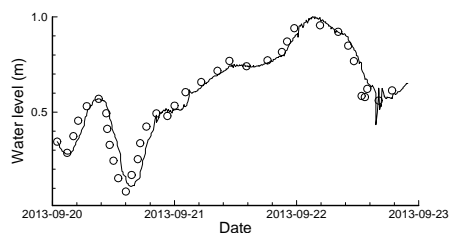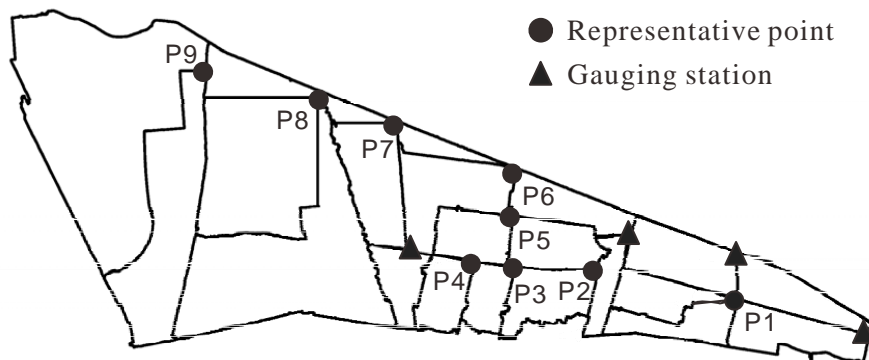
(a)

(b)

(c)

(d)

Figure 3 Comparison of validated water-level prediction (3-hr lead time) and measured data recorded during typhoon Usagi: (a) Zhongnanxing station (b) Xinnan station (c) Sijie station (d) Meifu station

19

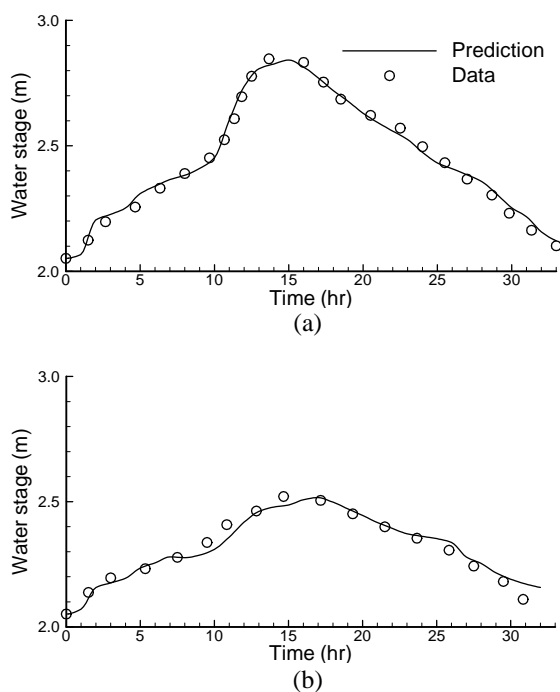**Figure 4 Divisions of sub-areas and representative points in Xinnan area**

(a)

(b)

Figure 5 Comparison of model predictions and IPDB data at representative point P8: (a) 450 mm daily rainfall used in calibration; (b) storms with 20 yr return period used in validation (3-hr lead time)
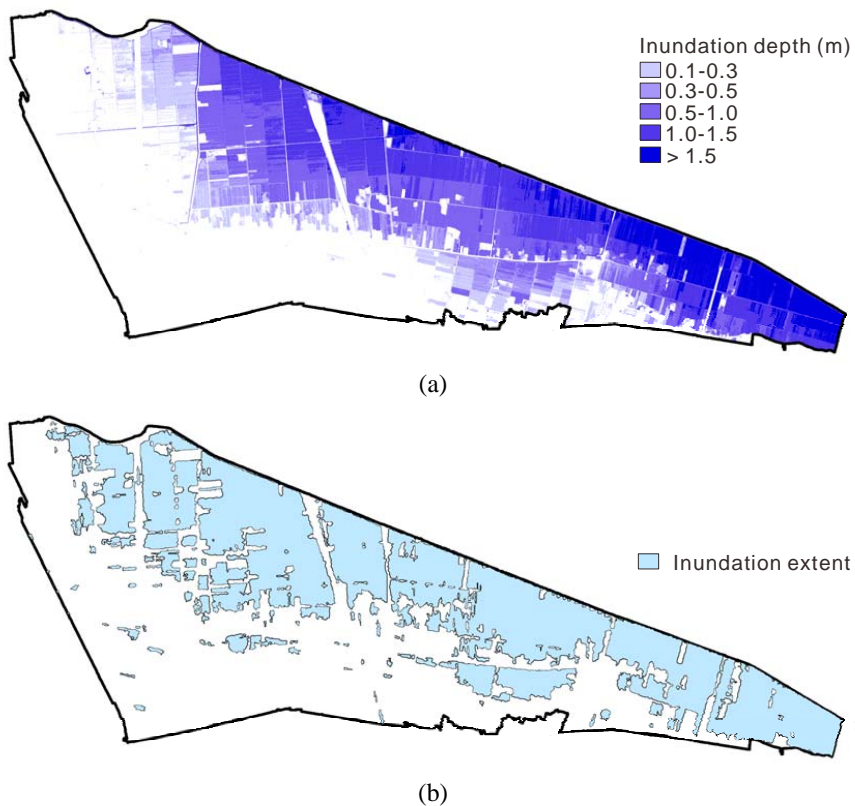
Figure 6 Comparison showing extent of inundation during typhoon Salola (2012/8/2 PM17:50): (a) SNTIX forecast (b) TerraSAR-X (courtesy of National Central University in Taiwan)