Inferring Digital Flood Mapping from Known Free Data

AMINUDDIN AB. GHANI, Professor & Deputy Director, River Engineering and Urban Drainage Research Centre (REDAC), Universiti Sains Malaysia, Engineering Campus, Seri Ampangan, 14300 Nibong Tebal, Penang, Malaysia. Email: redac02@eng.usm.my

CHUN KIAT CHANG, Science Officer, REDAC, Universiti Sains Malaysia, Engineering Campus, Seri Ampangan, 14300 Nibong Tebal, Penang, Malaysia. Email: redac10@eng.usm.my

CHENG SIANG LEOW, Research Officer, REDAC, Universiti Sains Malaysia, Engineering Campus, Seri Ampangan, 14300 Nibong Tebal, Penang, Malaysia. Email: redac21@eng.usm.my

NOR AZAZI ZAKARIA, Professor & Director, REDAC, Universiti Sains Malaysia, Engineering Campus, Seri Ampangan, 14300 Nibong Tebal, Penang, Malaysia. Email: redac01@eng.usm.my

ABSTRACT

Visits to areas in the subcatchments of Sungai Pahang were made throughout the year 2009 including Mentakab, Temerloh, Maran and Pekan to obtain information such as inundation areas, flood depth, flow discharge and water levels relevant to 2007 flood. Interviews with local people were also made to ascertain the true picture of the 2007 flood. Analyses of rainfall data, water level, flow discharge and river cross-sectional changes were carried out from data gathering. The development of Digital Elevation Model (DEM) was made from a combination of digital topographic maps and satellite images which were purchased from Jabatan Ukur dan Pemetaan Malaysia (JUPEM). This visually enhanced DEM is then used for the development of the actual extend of 2007 flood in the Study Area. Several options of flood mitigation works are proposed to reduce the impact of similar flood in the future based on the developed DEM.

Keywords: Sediment transport; alluvial river; flood mitigation; erosion; deposition.

1 Introduction

Malaysia is fortunate in that historically it has not experienced natural disasters in the form of earthquakes, volcanoes, and typhoons. The most common natural disaster frequently encountered in Malaysia is flooding. Two major types of floods occur in Malaysia, namely monsoon floods and flash floods. The Department of Irrigation and Drainage (DID) in Malaysia has estimated that about 29,000 km², or 9%, of the total land area and more than 4.82 million people (i.e. 22% of the population) are affected by flooding annually. The damage caused by flooding is estimated to be about RM 915 million (Chan, 2005).

Whilst monsoon floods are governed by heavy and long durations of rainfall, more localized flooding, which occurs especially in newly developed town areas, has been reported more frequently in recent years. In October 2003 major flooding affected a large area in the northwestern part of the Peninsular, including the states of Kedah, Penang and Northern Perak. The December 2007 flood (Figure 1), on the other hand, occurred in the state of Pahang, after more than 30 years (DID, 1974) since the last similar magnitude flood of 1971 (Figure 2, Tables 1 and 2). Flash floods have occurred more frequently in the country since the 1980s, with this type of floods often having a drastic impact on parts of the country.

Several major floods have been experienced in Malaysia for the last few decades. Flood occurrences seem to be getting more frequent in recent years, especially in some cities like Kuala Lumpur, Penang and Kuching where rapid urbanisation is taking place. The business of flood control and mitigation seems to be closely associated with development of Malaysia. After several dramatic flooding events struck the country, causing substantial lives and property losses since the 1960s, the government had taken several positive steps and seriously planning to envisage flood mitigation projects in its national plans, translated substantially by the establishment of the Natural Disaster Relief Committee in 1972 and the Permanent Flood Control Commission in December 21, 1971, specifically to study short-term measures to prevent the occurrence of floods and long-term measures for flood mitigation. Both structural and non-structural flood mitigation measures have been developed by the DID to be implemented for 17 major river basins and 27 towns nationwide including Sungai Pahang basin. The objective of river basin studies is to draw up appropriate flood maps and also feasible projects for the respective basin areas so that their development is properly managed and that water resources management, including flood control measures,

is effective and well-controlled. These studies recommend the optional flood control planning and design criteria for the respective basins. Generally, socio–economic considerations for the basin will dominate the design criteria (Chia, 2004).

In recent years, DID is more conscious of the need to carry out flood mitigation projects on a river basin basis rather than on a piecemeal basis. This kind of approach will involve a shift from the traditional thinking in terms of controlling flooding through expensive engineering structures to the more comprehensive approach of viewing the solution in terms of managing flooding by incorporating structural as well as non-structural measures. In this paper, results are presented, to formulate the study objectives, i.e. digitally map the 2007 flood inundation areas along Sungai Pahang by gather hydraulic and hydrologic data. In this paper, a digitally map of the 2007 flood inundation areas along Sungai Pahang was developed by gathering available hydraulic and hydrologic data.

2 Study Area

The Sungai Pahang basin (Figure 1) is located in the eastern part of Peninsular Malaysia and drains an area of 29,300 km² of which 27,000 km² lies within Pahang (about 75% of the State) and 2,300 km² is located in Negeri Sembilan. Sungai Pahang is the longest river in the Peninsular Malaysia at about 435 km. Sungai Pahang originates from Kuala Tembeling at the confluence of

two equally large and long rivers, about 304 km from the river mouth in the central north, the Sungai Jelai emerges from the Titiwangsa Range at the northwestern tip of the Sungai Pahang Basin, while the Sungai Tembeling originates from the Timur Range at the northeastern edge of the basin. Other main tributaries of Sungai Pahang are the Sungai Semantan, Sungai Teriang, Sungai Bera and Sungai Lepar.

The Sungai Pahang begins to flow in the south east and south directions from the north passing along several major towns such as Kuala Lipis, at the mouth of the river bearing the same name on Sungai Jelai; Jerantut, the gateway to Taman Negara Sungai Tembeling; and Temerloh, midway on the river at its confluence with Sungai Semantan and finally turning eastward at Mengkarak in the central south flowing through the royal town of Bandar Diraja Pekan near the coast before discharging into the South China Sea.

Sungai Pahang basin has an annual rainfall of about 2,136 mm, a large proportion of which is brought by the North-East Monsoon between mid October and mid January. Due to gentle terrains, the velocity drops and the river channel of the Sungai Pahang proper is wider and shallower compared to the Sungai Jelai and the Sungai Tembeling. The mean flow of Sungai Pahang measured at Station 3527410 (Lubok Paku) which is the most downstream stage station in the Sungai Pahang basin is 689 m³/s based on the latest 10-year (1999-2008) data (JICA, 2010). Detail of Sungai Pahang basin is shown in Table 1.



Figure 1 Sungai Pahang Basin Delineation

	Table 1 Sungai Panang Basin		
Location	Central part of Peninsular Malaysia, being sandwiched by the Titiwangsa Range in the west and the Timur Range in the east both of which are expanding in the north-northeastern direction.		
Area	29,300 km ²		
Length	435 km		
River bed slope	Sungai Pahang: 0.016% (1/6,200)		
	Sungai Jerai: 0.034% (1/2,900)		
	Sungai Tembeling: 0.024% (1/4,100)		
Population	1,190,000 (2010)		
Geographical regions	Tributaries: Tembeling, Jelai, Lipis, Tekman, Kerau, Semantan, Triang, Bera, Jengka, Jempul, Luit, Mentiga, Lenar and Serting		
	Mein lake, Lepar and Serting		
Land use	virgin jungle, rubber, paddy, oil palm, other agricultural crops, urban		
Ecosystems	Lowland tropical rainforests:		
	• Lowland Dipterocarp forest: Usually dense, with many thousands species of trees as well as shrubs herbs and woody climbers		
	• Hill Dipterocarp forest: Similar to lowland Dipterocarp forest		
	• Upper Dipterocarp forest: Characterised by Shorea platyclados		
	Lower and upper montane forest: Fagaceae and Lauraceae, Preris ovalifolia, Rhododendron spp. and		
	Vaccinium spp.		
	Peat swamp forests and mangrove forests: along the coast		
	Forest plantation: Pinus caribaea, Araucaria spp., Acacia mangium, Gmelina arborea and		
	Paraserianthes falcataria		
Climate	Daily minimum and maximum temperatures are around 23°C and 32°C (Except the highland area)		
	Annual rainfall varies from approximately 1,700 to 2,800 mm within the basin (mean annual rainfall		
	obtained from 10 years rainfall data of 1999-2008 is 2,136 mm)		

Table 1 Sungai Pahang Basin

Source: Summarized from JICA (2010)

2.1 Historical Flood Conditions

Several major floods occurred in the last few decades in Sungai Pahang Basin, causing extensive damage and inconvenience to the community. According to historical records, the flood of 1926 that affected most parts of Peninsular Malaysia is supposedly the worst in living memory in Malaysia. However, official records are too insufficient to describe the condition of that flood in detail. In January 1971, a catastrophic flood again swept across many parts of the country and it was considered the second largest flood. Pahang was severely affected, suffering great economic losses in the form of properties and crops, with the inundation area of about 3000 km^2 , 150,000 evacuees and loss of 24 lives, which led experts to conclude that the estimated flood damage to be around 38 million US dollars including the intangible damages. The scale of 1971 flood is over the 100-year ARI based on the hydrological probability analysis using the mean 8-day rainfall records (DID, 1974).

The December 2007 flood can be recognized as the third largest flood for Sungai Pahang Basin in terms of basin mean 8-day rainfall. The water level exceeded the danger level at the Lubok Paku, Temerloh, Pekan stations. According to field survey, the inundation depth ranges from 1.0 to 2.0 m at Pekan Center and from 0.5 to 2.0 at the major towns in Temerloh and Maran districts. The long duration flood forced people to stay at designated evacuation centres in Rompin, Maran, Kuantan, Pekan, Raub, Bera, Jerantut, Bentong, Temerloh and Kuala Lipis districts for as long as 22 days,

a total of 8 casualties were also recorded in the state of Pahang, with one casualty at Rompin, two at Pekan, three at Temerloh and two at the Maran district respectively. The flood damages were estimated at RM 263 million by DID (DID, 2010). A flood map, which was delineated by DID using satellite images is shown in Figure 2.



Figure 2 Sungai Pahang Flood Map for December 2007 (after DID, 2007)

3 Study Approach and Model Development

Two common approaches adopted in reducing the impact of flood problems have been increasingly adopted in Malaysia and these include structural and non structural measures. Structural measures include such measures as river widening, deepening and straightening, with the aim being to reduce the magnitude of the flood, but at the same time this approach often transfers the flooding problem further downstream. For non structural measures, tools such as computer models can be used to quantify the effects of human interference to the river system. Such tools are already available widely used in many countries worldwide, but the application of sophisticated models is still relatively new in Malavsia (Chang et al., 2008; Leow et al., 2009; Ab. Ghani, et al., 2010). One reason for this limited use of such models in Malaysia is that the tools often do not properly model the more extreme flood events, where the river flows are often supercritical. In Malaysia it is regarded as increasingly important to carry out a thorough analysis of flood events with the help of available river models to understand the flood behaviour before any structural measures are undertaken. Therefore, before any amendments are implemented within a river basin, river engineers must evaluate the potential extent and impact of flood events and advise the implementing agencies as to what steps shall to be undertaken to provide further preventative measures to avoid the anticipated flood problems that might occur (Ab. Ghani et al., 2009).

3.1 Flood Information Gathering

The types of data used in this study for building the model of the 2007 flood event are: digital topographical map, satellite images, land use, soil, ground model (digital elevation model or DEM), hydraulic and hydrologic data (flow discharge hydrographs, water levels, and tidal data), flood inundation areas and field observations. Interviews with local people and on-site ground survey and validation (Figure 2) also made to ascertain the true picture of the 2007 flood.

3.2 Inundation Area Mapping

The mapping first involves the creation of terrain model. The NASA SRTM Digital Elevation Model (90m x 90m) was used to profile the terrain. The Satellite SPOT-5 images obtained would then be overlaid and tied down to the terrain to better describe the geographical features of the Study Area. Figure 3 presents the flowchart describing the procedures executed to produce the corrected DEM. As SRTM DEM comes with 10m absolute error (Rodriguez et al., 2006), it was not quite sufficient for flood mapping. To improve the DEM, more than 150 exact field elevation points were collected using Differential-GPS (DGPS) with vertical accuracy of less than 0.3m in a field survey exercise in May 2010. The points were used to compare against the DEM elevations of same locations. The discrepancies were averaged out and applied as correction to the DEM. This enhanced DEM was then used for flood mapping.



Figure 3 DEM Development

In order to profile the flood, maximum water levels at 4 gauge stations (namely, Temerloh, Mengkarak, Chenor and Triang) during the 2007 flood were used. Linear interpolation method was used to profile the water level gradient along the river. The result was converted into TIN and subsequently raster file. By overlaying the flood profile raster with DEM, areas where flood occurred (flood raster > DEM) can be identified. This areas are then reclassified and generalised to form the flood inundation map for the study area. The records of flood elevation obtained can then be profiled into the DEM to study the actual extend of flood in the Study Area. The GIS workflow to produce this flood inundation map is described in Figure 4 while Figure 5 shows the obtained DEM with preview of SPOT-5 satellite images overlaid in 2D mode.



Figure 2 On-Site Ground Survey and Validation using Professional Data Mapper for GPS Data Collection and Mapping



Figure 4 Flood Map Development

3.3 Hydraulic and Hydrological Data Analysis

Analysis of hydraulic and hydrological data will benefit and provide data for future hydraulic modelling works. Accurate information about the flow rates of rivers is important for a variety of hydrologic applications such as water and sediment bed material load estimation, water resource planning, operation and development, and hydraulic and hydrologic modelling (Guven and Aytek, 2009). However, collecting data for discharge on a continuous basis is costly, especially during large flood events. An alternative approach is to convert records of water stages into discharges using a stage-discharge relationship. In the present study, a mathematical models for the estimation of stage-discharge relationships based on the Gene-Expression Programming (GEP) techniques has been carried out. GEP is a new evolutionary Artificial Intelligence technique developed by Ferreira



Figure 5 DEM Development

(2001). This technique is an extension of GP, developed by Koza (1992).

The daily discharge data at Station 3424411 (Temerloh) for the year 2007 were chosen for the training of the proposed GEP models, and the 2004 daily discharge data were chosen for the testing of the models. The 2007 daily discharge data were chosen for training because the worst recent flood occurred in that year and the range of discharges also gave a wider spectrum. Training with a wider data spectrum can ensure a robust model to predict discharges over a wider range and better estimation of maximum flooding in extreme events (Azamathulla et al., 2011). The discharge (Q) is modelled in terms of the stage (S) using a GEP approach. Initially, the "training set" is selected from the entire data set, and the rest is used as the "testing set". Once the training set is selected, one could say that the learning environment of the system is defined. Detailed modelling

which includes five major steps to prepare the GEP for use is well described in Azamathulla et al. (2011). The explicit formulations of the GEP for discharge as a function of stage were obtained in simplified form as follow:

$$Q = 9.84S^2 - 64.391S - 4033.296 \tag{1}$$

Figure 6 shows the expression tree (ET) for the above formulation.



Figure 6 Expression Tree for the GEP Formulation for the Temerloh Station, where d0=S

4 Results and Discussion

4.1 Digital Flood Map

A comparison has been made between simulated flood map with the DID flood map (delineated using satellite images) in order to verify the performance of the SRTM dataset for flood inundation mapping. To achieve this, an indicator proposed by Horitt (2006) was used. Equation 2 gives the mathematical expression of the indicator.

$$F = \frac{A - B}{A + B + C} \tag{2}$$

Where F is the goodness-of-fit, A is the wet area correctly predicted by simulation, B is the wet area wrongly predicted as dry (under prediction), C is the dry area wrongly predicted as wet (over-prediction). An Fvalue of near to 1 indicates model flood extent prediction is close to the observed flood inundation area.

An assessment using the indicator given by Equation 2 were made for flood in the region between Mentakab and Chenor. Table 2 sums up the calculations and results. The areas required were derived using GIS. The results are also mapped on the same sheet for comparison, i.e. Figure 7.

Table 2 Verification of Flood Inundation Map		
Parameters	Area (km ²)	
Total Test Area	1,411.06	
Predicted Flood	104.79	
Actual Flood	149.42	
Α	41.55	
В	107.86	
С	63.24	
F	-0.312	

It can be observed that the result for flood map verification is not encouraging. With F = -0.312, it can be said that the produced flood map are not in good agreement with the actual flood condition. This poor result can be attributed to:

- Accuracy of the DEM. While effort has been done to apply correction, it is perhaps not sufficient. On top of that the correction applied also contains uncertainties. The current elevation correction is very crude, i.e. there is no consideration for vegetation/ building removal prior to correction. This will affect the DEM.
- Lack of water level station, especially for the tributaries. In Figure 7, a large flood area was missed by the model, as it is beyond the most upstream gauge (Temerloh). The linear extrapolation method used in this study is clearly not sufficient to profile flood coming into the Pahang River from tributaries.
- Uncertainties in flood delineation from satellite. While it is the most convenient way to derive flood map for large area, certain uncertainties do involved. For example, floods in tropical climate occur in monsoon with heavy clouds cover, which require cloud removal that induces uncertainty to the map. One example would be Point 1 (Figure 7) where flood are mapped on ground with elevation much higher than the flood record.

In response to these weaknesses, the authors are currently improving the methodology with the following enhancement:

- Further site survey to collect more DGPS points are currently undergoing. With a larger sample pool that covers a wider area, the reliability of the correction to be applied can be increased.
- Developing a mechanism to remove building/vegetation height from DEM to create Digital Terrain Model (DTM) which will better represent the terrain of the study. Instead of regional correction applied, corrections will be determined based on land use type, where presumably, each land use would have a certain average height of vegetation or building to be removed and corrections to be applied.



Figure 7 Comparison of Flood Inundation Area of Model with Actual Condition

4.2 Hydraulic and Hydrological Data Analysis

Different input sets of daily stage-discharge records were developed to calibrate the GEP models. The GEP estimates were then compared to the observed data via scatter plots for the training (Figure 8a) and testing sets (Figure 8b). As shown in figures, it was evident that the proposed GEP model determined the nonlinear relationship between the input and the output variables impressively well, with an R² of 0.993 and an RMSE of 62.34 m³/s. Comparing the GEP predictions with the observed data for the test stage demonstrated a high generalization capacity with R² = 0.945 and RMSE = 78.98 m³/s for the 2004 data. Therefore, the modelling demonstrated the acceptable performance of the GEP models for estimating discharge in both the training and testing stages.

4.3 Proposed Structural and Non-Structural Works

It was observed that flood in the river basin is extensive starting from Temerloh. It can be argued that excess flow from headwaters could not be contained within the river channel starting from the said area. As water overspills the river banks, flood occurs in a very extensive manner, submerging huge area of land, including riverside towns.



Figure 8 Scatter Plots in the GEP-based Training Period for the Temerloh Station (Azamathulla et al., 2011)

These observations have led to the proposed alternatives of a flood bypass channel and a stretch of river bund along the channel (Figure 9). First a flood bypass channel was proposed to provide alternative route for flood water right before it enters Temerloh. The channel will directly convey flood water further downstream, limiting flood damage in Mentakab and Temerloh. As flow is redirected, it is estimated that flood level would drop. However, critical areas such as Temerloh and Chenor townships would require bund constructions along river banks. The height and length of bund would require further study. In the next phase of the study, hydrodynamic model would be used to gauge the performance of these flood mitigation alternatives. First, the DEM would be modified to include the bypass channel and river bunds. Hydrodynamic model simulation based on the 2007 flood would be used as hydrological input. The reduction in flood depth and extent would be used to gauge the performance of the proposed alternatives. The hydrological and hydraulic regime and response of the entire river system will also be investigated to ensure sustainability of the proposed alternatives.



Figure 9 Proposed Flood Mitigation Alternatives for Sungai Pahang

5 Conclusions

A thorough analysis of flood events with the help of available river models to understand the flood behaviour before any structural measures are undertaken is a better option before implementing highly expensive flood mitigation projects. DEM development for a major river basin such as Sungai Pahang is deemed very important before any river modelling is applied. As such this study has developed the required DEM for Sungai Pahang utilizing the 2007 flood data and reconnaissance data from field survey, albeit some further improvement is required. It is hopeful that the final DEM will help improve the river modelling results for better planning and flood mitigation options.

Gene Expression Programming (GEP) also has been use as an alternative tools for modelling the stagedischarge relationship for the Sungai Pahang in the study. The RMSE and R^2 values were calculated to measure the deviation from and approximation of observed flows obtained from the Temerloh station. The overall results produced R^2 values very close to 1, suggesting very little discrepancy between observed and predicted discharges. Besides that, the RMSE remained at a very low level also confirm the use of GEP as an effective tool for forecasting and the estimation of daily discharge data in flood event.

Acknowledgments

The authors wish to express their sincere gratitude to Universiti Sains Malaysia for funding a research university grant (1001.PREDAC.814056) to conduct this research. Appreciation also goes to Department of Irrigation and Drainage Malaysia for providing data, as well as staff of REDAC, Universiti Sains Malaysia, for their involvement in the research work.

References

 Ab. Ghani, A., Ali, R., Zakaria, N.A., Abu Hasan, Z., Chang, C.K. & Ahamad, M.S.S. (2010). A Temporal Change Study of the Muda River System over 22 Years. *International Journal of River Basin Management*, IAHR, IAHS & INBO, Vol. 8, No. 1, pp. 25-37.

- Ab. Ghani, A., Zakaria, N.A. & Falconer, R.A. (2009). Editorial, River Modelling and Flood Mitigation: Malaysian Perspectives, *Water Management Journal*, Vol. 162, No.1, pp. 1-2.
- Azamathulla, H.Md., Ab. Ghani, A., Leow, C.S., Chang, C.K. & Zakaria, N.A. (2011). Gene-Expression Programming for the Development of a Stage-Discharge Curve of the Pahang River. *Journal of Water Resources Management*, Vol. 25, No. 11, pp. 2901-2916.
- Chang, C.K., Ab. Ghani, A., Abdullah, R. & Zakaria, N.A.(2008). Sediment Transport Modeling for Kulim River: A Case Study. *Journal* of Hydro-Environment Research, IAHR & KWRA, Vol. 2, No.1, pp. 47-59.
- Chia, C.W. (2004). Managing Flood Problems in Malaysia. *Bulletin Ingenieur*, Vol. 22, pp. 38-43. Board of Engineers Malaysia, Kuala Lumpur.
- Chan, N.W. (2005). Sustainable Management of Rivers in Malaysia: Involving All Stakeholders, *International Journal of River Basin Management*, Vol. 3, No. 3, pp. 147-162.
- 7. Department of Irrigation and Drainage or DID (1974). Pahang River Basin Study, Vol. 3: Basin Hydrology and River Behaviour.
- 8. Department of Irrigation and Drainage or DID (2007). *Annual Flooding Report of Pahang State* 2007. DID, Pahang.
- 9. Ferreira C. (2001). Gene Expression Programming in Problem Solving, 6th Online World Conference on Soft Computing in Industrial Applications (invited tutorial).

- Guven A and Aytek A. (2009). A New Approach for Stage-Discharge Relationship: Gene-Expression Programming, *Journal of Hydrologic Engineering*, Vol. 14, No. 8, pp. 812-820.
- 11. Horritt, M.S. (2006). A Methodology for the Validation of Uncertain Flood Inundation Models. *Journal of Hydrology*, Vol. 326, pp. 153-165.
- 12. Japan International Cooperation Agency or JICA. (2011). The Preparatory Survey for Integrated River Basin Management Incorporating Integrated Flood Management with Adaptation of Climate Change. *Draft Final Report: Volume 3 Pahang River Basin.* JICA, Tokyo.
- 13. Koza JR. (1992). *Genetic Programming: On the Programming of Computers by means of Natural Selection.* The MIT Press, Cambridge, MA.
- Leow, C.S., Abdullah R., Zakaria, N.A., Ab. Ghani, A. & Chang, C.K. (2009). Modelling Urban River Catchment: A Case Study in Malaysia. *Proceedings of the Institution of Civil Engineers: Water Management*, Vol.162, No. 1, pp. 25-34.
- Rodriguez, E., Morris, C.S., Belz, J.E., Charpin, E.C., Martin, J.M., Daffer, W., and Hensley, S. (2006). An Assessment of the SRTM Topographic Products. Jet Propulsion Laboratory, NASA, US. [online] Available at: http://www2.jpl.nasa.gov/ srtm/SRTM_D31639.pdf [Accessed 08 March 2010].