Challenges on Modelling a Large River Basin with Scarce Data: A Case Study of the Indus Upper Catchment

A. Sugiura [1,*], S. Fujioka [2], S. Nabesaka [2], T. Sayama [1], Y. Iwami [1], K. Fukami [3], S. Tanaka [1], K. Takeuchi[1]

[1] International Centre for Water Hazard and Risk Management, PWRI, 1-6, Minamihara, Tsukuba, Ibaraki 305-8516, Japan

[2] Japan Water Agency, Japan

[3] National Institute for Land and Infrastructure Management, Ministry of Land, Infrastructure, Transport and Tourism, Japan

Peer Review History[1]

Abstract: *The unprecedented floods of 2010 in Pakistan highlighted the necessity of a well-calibrated hydrological model of the Indus upper catchment for a comprehensive flood risk assessment. However, this modelling was an extremely challenging exercise because of the lack of hydrometeorological data, which are difficult to collect due to the geography of the catchment. In the study area (133,300 km^2), there are 24 raingauges collecting sufficient daily data, which leads to an average area of Thiessen polygons well over (by 10 times) the WMO minimum density network requirements of 250 km^2 for hilly area. The lack of local data for soil and aquifer poses another challenge. Despite those limitations, IFAS (Integrated Flood Analysis System) was run to conduct rainfall runoff analysis from the very upstream (in India and China) to Taunsa (midstream Indus in Pakistan). The 30 sec Digital Elevation Model based on GlobalMap Elevation from ISCGM was upscaled to a 5 km grid model. The runoff analysis engine of IFAS was based on a 3-layered spatially distributed tank model. The daily precipitation data from 24 raingauges, discharge data from 9 river stations, barrages and dams and NCEP reanalyzed latent heat fluxes were considered as input data. Global datasets for land cover (Global Map Land Cover, ISCGM) and soil textural types and depths (FAO/UNESCO DHSM) were used for parameterization. The upper catchment was divided into sub-basins and calibration was conducted independently for each of them. As simulated discharges for mid-lower stream sub-basins were more reasonable than for more upstream sub-basins, parameters calibrated in the mid-lower sub-basins were applied to the upstream ones. Then, the calibration was conducted for three flood events (1988, 1997 and 2010). Finally, in order to validate the parameters and the model, Nash-Sutcliffe efficiencies (ENS) were calculated for discharges simulated for three other flood events (1992, 1994 and 2012). In average, ENS values were found over 0.80 at seven river stations and the model was considered well calibrated.*

Keywords: *Flood, hydrological modelling, large river basin, Indus, Pakistan, IFAS.*

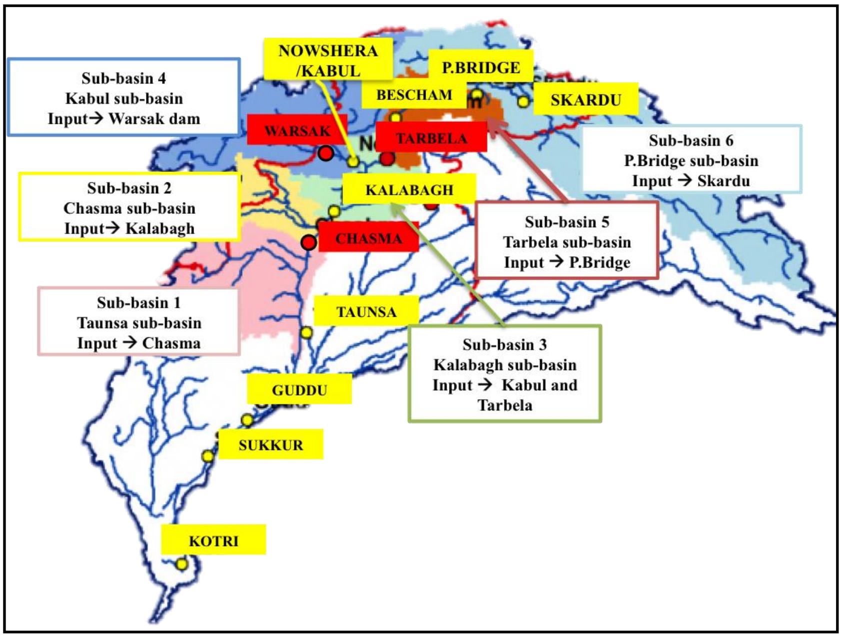
1. Introduction

Flood occurs frequently in Indus River Basin due to heavy rainfall during the monsoon season (July-September), exacerbated sometimes with increased snowmelt contribution to discharge (Inam et al., 2007; FFC, 2013). After 2010 unprecedented floods in Pakistan, Pakistani authorities highlighted the need to develop improved flood forecasting models (FFC, 2013). This research is part of the "Strategic Strengthening Flood Forecasting and Management Capacity in Pakistan" implemented by UNESCO from January 2012, and the final goal of this research is the development of a flood forecasting system. This paper focuses on the development of a hydrological model using the physically distributed rainfall-runoff model (Public Work Research Institute – Distributed Hydrological Model, 3-layer tank PWRI-DHM) mounted in IFAS (Integrated Flood Analysis System) on mainstream Indus from its source to mid-stream Taunsa (133,300 km^2, 37% of the

[1] *Paper JHER0203, submitted on 31/10/2014, accepted for publication after peer review and subsequent revisions on 22/12/2014*

* *Corresponding author may be contacted at aisugiura@gmail.com*

whole river basin, Figure 1a) and the verification of its reproducibility for 6 past flood events. Based on the lack of data availability and on the day order of magnitude lead-time between the different stations of the basin at which forecast and flood alert are issued, the study area was divided into 6 sub-basins and the model was calibrated separately using discharges of upstream stations as boundary condition. The objective of this paper is to present the methodology adopted to develop a calibrated hydrological model and verify its reproducibility despite the lack of data in the Indus river basin. Snowmelt contribution to discharge is only considered indirectly by giving boundary condition since snowmelt component modelling is out of the scope of this paper.



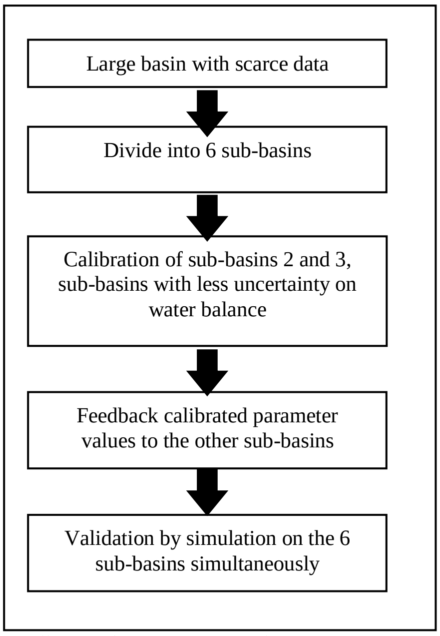


Figure 1a (left) Division of Upper-Indus into 6 sub-basins (river stations are shown as yellow and reservoirs as red).

Figure 1 b (right) Flow chart illustrates methodology.

2. Data availability in the target area

The Indus River system takes its source at 5182 m in Tibet (Inam et al., 2007) and crosses India, Afghanistan to reach the Arabian Sea in Pakistan where it mostly lies (FAO, 2009). The Indus River basin comprises the Western rivers, Indus mainstream with its tributaries like Shyok, Shigar, Gilgit and Kabul and Jhelum, and the Eastern Rivers, Chenab, Ravi, Beas and Sutlej (Inam et al., 2007). This section describes the data availability for the study. Following a consistency test comparing measured rainfall data from different sources in Pakistan, Pakistan Meteorological Department (PMD) daily rainfall data were selected as the most reliable source of measured rainfall. Hence, PMD daily rainfall data collected from 03:00 to 03:00 GTM were the only complete source available and were therefore chosen as the input data for both calibration and validation processes. Figure 2 describes the distribution of PMD raingauges and meteorological stations over Pakistan. The number of stations for which data are available varies from year to year, illustrating the effort of PMD to develop an operational meteorological data network with the number of stations almost tripled from 36 to 92 between 1988 and 2012 over the whole river basin in Pakistan. Secondly, reference evapotranspiration ETo (FAO-Penman-Monteith (Allen et al., 1988)) can be calculated at 18 PMD meteorological stations in the whole Indus river basin. There are two areas with no measured evapotranspiration even after distributing ETo using Thiessen polygons (Figure 2, circles). Therefore, National Centres for Environmental Predictions (NCEP) latent heat net fluxes, available globally, were chosen as an estimator of evapotranspiration. NCEP latent heat net fluxes daily mean value (W/m^2) are available monthly, at 1.9 ° grid, from 88.542 N-88.542 S, 0 E-358.125 E and are calculated from 30 years data (1979-2009) (Kanamitsu et al., 2002). Thirdly, 6 hourly measured discharges provided by WAPDA (Water and Power Development Authority, Pakistan) at the following seven stations' discharges were considered: Partab Bridge, Besham, Nowshera, Tarbela, Kalabagh, Chasma and Taunsa (refer to discharge stations in Figure 2). In addition, discharges data for Skardu, the most upstream station and for Warsak dam, in the upstream of Nowshera station were also included as boundary conditions. Fourth, available soil hydraulic properties data do not cover the whole target area. Data collected so far are very localized e.g. Kelleners et al. (1999) study on soil hydraulic properties for unsaturated zone around Faisalabad in Punjab province.

In summary, both local PMD evapotranspiration and soil hydraulic properties data were found to be very limited and it was necessary to rely on global datasets and PMD daily rainfall data distributed using Thiessen polygons, NCEP reanalysis data for latent heat net fluxes and FAO/UNESCO globally available soil type distribution map (FAO, 2009) were selected as default input data.

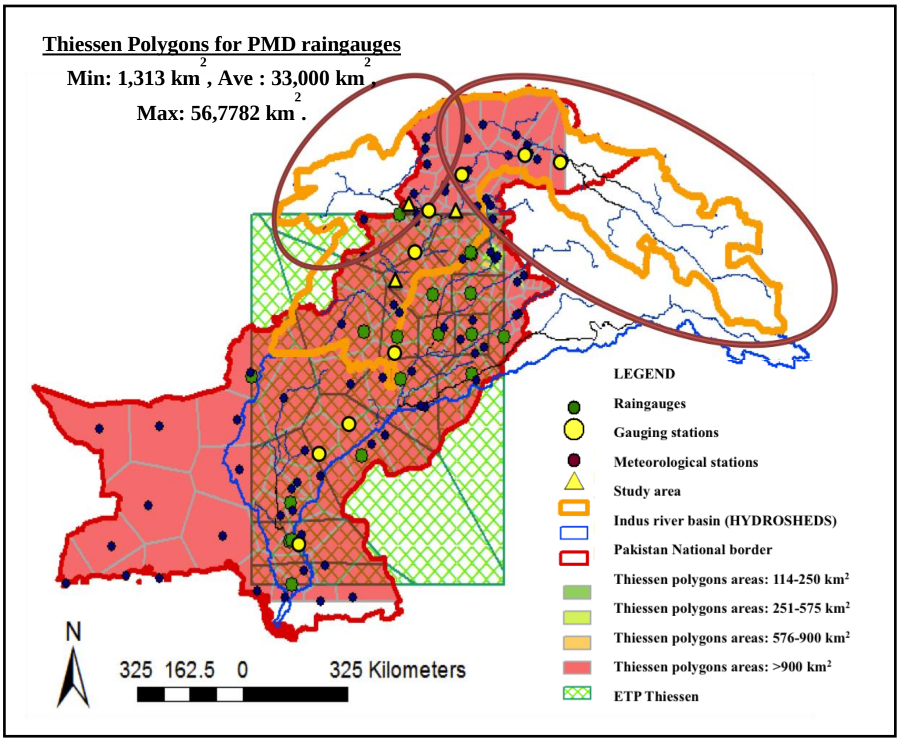


Figure 2 Raingauges, meteorological stations, discharge stations and reservoirs distribution. Thiessen polygons areas are in km^2 (mostly red or over 900 km^2 for raingauges, green dash for ETP). Circles indicate territories without any coverage of raingauges: Kabul river basin and very upstream of the Indus.

Table 1 Description of the 3-layer model PWRI-DHM

Model	Function
Surface tank model	Infiltration to unsaturated layer, surface runoff, surface storage, evapotranspiration, rapid unsaturated subsurface flow
Subsurface tank model (for the 3 tank model)	Infiltration to aquifer, subsurface runoff, subsurface storage, slow unsaturated subsurface flow
Aquifer tank model	Outflow from aquifer, aquifer loss
River tank model	River course discharge

3. Methodology

3.1 PWRI-DHM in IFAS 3-layer tank model parameterization

PWRI-DHM is based on Sugawara et al. (1956) tank model concept, used in a distributed and 3-layer tank configuration. The 3 layers are (Figure 3 and Table 1): firstly, a surface layer tank presenting a set of parameters replicated into land use classes based on GlobalMap Land cover (ISCGM) 30 sec resolution) categories grouped into five classes of parameters according to land use (i) forest and woodland (ii) shrubs, herbaceous or bare land (iii) cropland, paddy field, and wetland (iv) urban and (v) snow/ice and water bodies. Secondly, an unsaturated layer tank and an aquifer layer tank presenting a set of parameters replicated into soil textural classes, which flows directly or indirectly to the river routing model in the river tank presenting a set of 10 parameters, with in total 130 parameters to tune. Details are available in Sugiura et al. (2010) for IFAS, and Fujita et al. (2006) for the 3-layer tank model. In this study, a 5-km mesh model was set up. The parameterization was performed by trial and error for the surface tank by essentially tuning infiltration capacity, maximum storage height, surface roughness coefficient, for the aquifer tank by essentially tuning slow intermediate flow regulation coefficient and baseflow coefficient and river tank by essentially tuning Manning's roughness coefficient, those parameters being the most sensitive. For the unsaturated tank, all parameters were fixed according to Maidment (1993) hydraulic properties corresponding to soil textural classes. Moreover, for the most northern area of the Indus river basin, two uncertainties had to be overcome, firstly the lack of measured rainfall data and secondly the fact that precipitations for this high elevation part (over 7000 m elevation for some parts) are snowfall and not rainfall. But snowfall estimates and snowmelt modelling are not within the scope of this study. Hence, the solution retained was first to divide the study area into six sub-basins (Figure 1a). Then, for each sub-basin, upstream station discharges were given as the boundary conditions and simulated discharges were compared to measured discharge at the outlet (downstream) of each sub-basin for calibration and validation. By giving measured discharge as boundary conditions, it was expected to account indirectly for snowmelt runoff. Hence, at Skardu, Partab Bridge and Nowshera river stations, discharges were taken as input to compensate the lack of rain/snowfall and snowmelt data. To take into account Warsak, Tarbela dams and Kalabagh, Chashma, Taunsa barrages operation, outflows measured at those points were also given as boundary conditions. Therefore, the discharges between the following points were simulated: between Skardu and Partab Bridge (Sub-basin 6), Partab Bridge and Tarbela (Sub-basin 5), Tarbela and Kabul (Sub-basin 4), Kabul, Tarbela and Kalabagh (Sub-basin 3), Kalabagh and Chashma (Sub-basin 2), then Chashma and Taunsa (Sub-basin 1).

3.2 Characteristics of the six sub-basins

The study area covered 133,300 km^2 and was divided into 6 sub-basins according to the following characteristics (Figure 1a). Sub-basins 1 and 2 characteristics are the absence of raingauges on part of their area and the limited number of meteorological stations. Sub-basin 3 is characterized by a greater but yet insufficient number of raingauges (minimum of 1812 km^2 per station, maximum of 9958 km^2 per station and an average of 4791 km^2 per station), all over the recommended minimum number of non-recording rain gauges network of 250 km^2 for mountainous area (WMO, 2008). Sub-basins 4, 5 and 6 (Figure 2, circles) characteristics are the absence of raingauges on most of their areas, the absence of meteorological station, and a significant contribution of snowmelt in their discharges (Inam et al., 2007).

Because of the principle of equifinality, it is always possible to find different sets of parameters giving good fits with observed data (Beven and Freer, 2001). However, if the objective is to keep the model somehow with its physically distributed characteristics, calibration should take place where uncertainty is lower. In our case, Tarbela and Chasma dams present more reliable inflow and outflow data than at barrages. Moreover, the numbers of hydrometeorological data are more widely available in the mid-downstream part of the target area. Therefore, calibration was performed on sub-basins 2 and 3 as explained in 3.1 and their tuned parameter values were then fed back to the other sub-basins. Moreover, the aim of this modeling is to use the model as part of a flood forecasting system; therefore, the calibration needs to be performed for different magnitude of floods. Therefore, considering also the data availability and because 1998, 1992, 1994, 2010, 2011 and 2012 were severe flood years (FFC, 2013), with 2010 being an extreme flood, the three flood events (1988, 1997, 2010) were considered for calibration, and three others for validation (1992, 1994, 2012) .

Moreover, because lead times in the Indus river basin can be counted in days, for instance lead time of 24 hours between P. Bridge and Tarbela, 26 hours between Tarbela and Kalabagh, 51-72 hours between Chasma and Taunsa; this configuration using upstream sub-basin measured discharge as boundary condition to minimize uncertainty on a given sub-basin should allow forecasting discharges downstream.

4. Results

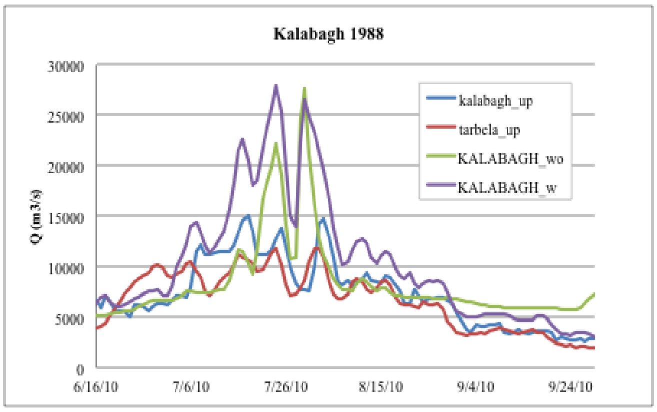
Nash-Sutcliffe Efficiency (ENS) (Nash and Sutcliffe, 1970) was selected to assess the performance of IFAS by inputting boundary conditions to account for the contribution of upstream catchment feeding into each of the 6 sub-basins in the target area. The obtained ENS values are reported in Table 2.

ENS values were calculated for each of the calibration flood events and sub-basins. The average value for all the stations and events confounded is found to be -1.16 if no measured discharge data are given as boundary condition and 0.61 if they are given. Without boundary condition, ENS are mostly negative, indicating the mean value of observed data is performing better as a predictor than our model. However, if measured discharges are input as boundary condition, the performance of the model increases remarkably with ENS values over 0.9. The results show that PWRI-DHM is able to simulate rainfall-runoff processes provided sufficient data are given. We will now focus on the discussion of the results from the configuration with measured discharges input as boundary conditions to each sub-basin. For 1988, the performance of the model is low in general (ENS even negative for Kalabagh). The limited number of raingauges for 1988 may explain these poor results. Hence, rainfall input being insufficient, the simulated discharges are grossly inaccurate. Moreover, for Kalabagh, ENS is low (0.45) for 1988 and better for 1997 and 2010 (0.85 and 0.83, respectively). Figure 4 compares hydrographs at Kalabagh obtained from simulation with or without boundary condition given in Tarbela. It appears that the two excessive peaks in the end of July 1988 are mainly due to uncertainties on rainfall data for sub-basin 3 for 1988.

Table 2 Nash-Sutcliffe Efficiency (ENS) in calibration and validation, firstly with discharges given as boundary condition (white background) and secondly without boundary condition (grey background). (Score under 0.5 are presented in red)

	Calibration: average ENS = 0.61 with discharge input, ENS = -1.16 without discharge input.						Validation: average ENS=0.67 with discharge input, ENS = -1.60 without discharge input.					
	1988		1997		2010		1992		1994		2012	
Taunsa	0.89	-0.03	0.86	-2.40	0.96	-0.03	0.85	-1.46	0.93	0.55	0.10	-16.39
Chasma	0.82	-0.29	0.86	-0.61	0.95	0.03	0.92	-1.76	0.93	0.65	0.95	-6.15
Kalabagh	-0.27	-0.34	0.85	0.01	0.83	-0.03	0.91	-1.80	0.86	0.63	0.89	-4.20
Kabul	0.34	0.73	0.69	0.69	0.28	0.48	NoData	NoData	0.75	-0.57	0.71	-4.22
Tarbela	0.38	-0.45	0.78	-0.18	0.73	-0.59	0.8	-3.95	0.92	-0.23	0.92	0.06
Besham	0.2	-0.62	0.79	-0.25	0.76	-0.67	0.72	-4.49	0.9	-0.41	0.93	-0.05
P.Bridge	0.71	-0.25	0.85	-0.06	0.52	-1.18	0.56	-6.53	-1.18	-0.38	0.02	-0.02

For Kabul, ENS values are lower than for other stations, with an average of 0.56. This is mainly due to uncertainties on rainfall data. Indeed, as reported in section 2, over the Kabul river basin, there is only one rain gauge station covering an area over 52,636 km^2. Therefore, measured discharges at Warsak dam were taken as boundary conditions. We did not try to fine-tune further as it would hide the model performance due to the lack/uncertainties associated with the rainfall data. And for 2010, the raingauge was washed away during the highest stage of the flood (PMD personal communication) and therefore, there is no data to compare the simulation with. The calibration results are deemed satisfactory and we will now consider the performance of the model for the flood events 1992, 1994 and 2012 to validate the calibration process. For part from P. Bridge, where snowmelt contribution is not negligible, ENS values are satisfactory (average of 0.82 without P. Bridge). The model manages to simulate properly the trends (increase or decrease) according to rainfall input. In particular, peaks timing and intensities are correct. For 2012, the poor performance of the model for ENS for Taunsa was unexpectedly low (0.10). However, after comparing the discharges at Taunsa for 1992, 1994 and 2012 (Figure 5), it appears that in 2012, discharges were significantly lower (almost 50% lower) than those during the other years. Moreover, the model response to rainfall is appropriately simulated and the trends are properly reproduced. This points out the strong dependency of the model to the availability and quality of measured discharge data.



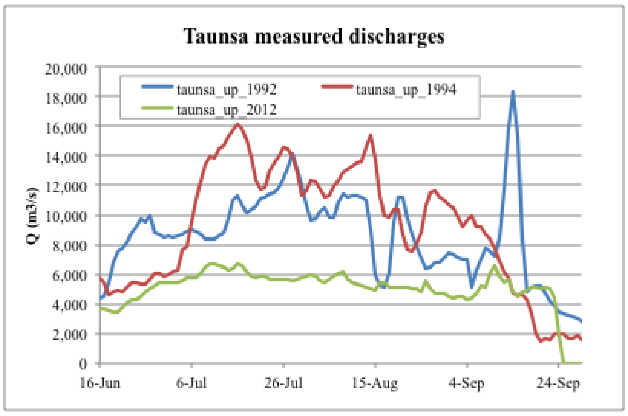


Figure 4 (left) Comparison between simulation of discharges in Kalabagh with (KALABAGH_w) or without (KALABAGH_wo) input of Tarbela discharges (tarbela_up) against observed data (kalabagh_up) for 1988.

Figure 5 (right) Comparison of observed discharges in Taunsa for the year 1992, 1994 and 2012 from the 15th of June until 1st of October of each year.

5. Conclusion

This study presents the calibration and validation of a distributed physically based rainfall-runoff model known as PWRI-DHM for the Indus river catchment in Pakistan. It has been found that the model performance is strongly dependent on the availability and quality of measured discharge and precipitation data. Because the flow concentration times in Indus system are of 'days order of magnitude', dependence on measured discharge data may not impede the model use efficiency as a flood forecasting system. The findings of this paper will be useful to similar flood study projects in other countries.

6. Acknowledgements

This study is part of the UNESCO project "Strategic Strengthening of Flood Warning and Management Capacity of Pakistan" funded by JICA. The authors acknowledge and thank the cooperation of all the project partners, in particular Pakistan Meteorological Department for providing necessary data for the study.

References

Allen RL, Pereira, DR, Smith M (1988). Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. Paper 56, Irrigation and Draianage, FAO, Rome, Italy, 300pp.

Beven K, Freer J (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *Journal of Hydrology*, 249, 1-4, 11-29.

FAO/IIASA/ISRIC/ISSCAS/JRC (2009). Harmonized World Soil Database (version 1.1). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

FFC (2013). Annual Flood Report 2012. Federal Floods Commission (FFC).

Fujita K et al. (2006) Project Research Report 299: Watershed/Urban Regeneration in Accord with Nature Technical Report (II). River Environment Laboratory, NILIM, Tsukuba, Japan: NILIM.

Inam A, Clift PD, Giosan L, Tabrez AR, Tahir M, Rabbani MM, Danish M (2007). The Georgraphic, Geological and Oceanographic Setting of the Indus River. In Large Rivers: Geomorphology and Management, edited by A Gupta, John Wiley and Sons.

Kanamitsu M, Ebisuzaki W, Wollen S-K Y, Hnilo JJ, Fiorino M, Potter GL (2002). NCEP-DEO AMIP-II Reanalysis (R-2), *Bulletin of the American Meteorological Society*, 1631-1643.

Kelleners TJ, Beekma J, Chaudhry MR (1999). Spatially variable soil hydraulic properties for simulation of field-scale solute transport in the unsaturated zone, *Geoderma*, 92, 3-4, 199-215.

Maidment DR (1993). Handbook of Hydrology, McGraw-Hill.

Nash JE, Sutcliffe JV (1970). River flow forecasting through conceptual models, Part I - A discussion of principles, *Journal of Hydrology*, 10, 282–290.

Sugawara M, Maruyama F (1956). A method of prevision of the river discharge by means of a rainfall 555 models. Symposia Darcy. Dijon: International Association Science Hydrological Publication, 42, 3, 556, 71-76.

Sugiura T, Fukami K, Fujiwara N, Hamaguchi K, Nakamura S, Hironaka S, Nakamura K, Wada T, Ishikawa M, Shimizu T, Inomata H, Itou Z (2010). Experimental application of flood forecasteing system (IFAS) using satellite-based rainfall. 9th International Conference on Hydroingormatics. Tianjin, China: HIC2010, 1786-1793.

WMO (2008). Guide to Hydrological Practices, Sixth Edition, 168, WMO, Geneva.