Rainfall Runoff Modeling in Upper-Awash Sub-Basin

By Deksyos Tarekegn[1], Felix Mtalo[2] and Al-Weshah R.[3]

[1] Hydrologist, Ministry of Water Resources, Addis Ababa, Ethiopia
deksyost@yahoo.co.uk

[2]Department of Water Resources Engineering, University of Dar es Salaam, Tanzania.
mtalo@wrep.udsm.ac.tz

UNESCO Cairo Office, P.O. Box 11541, Cairo, Egypt, r.weshah@mail.unesco.org.eg

Abstract

Rainfall Runoff modeling is an important tool in the study of water resources and water management of the watersheds. Rainfall runoff models are mainly used for river flow forecasting for the management of the resource and to minimize the ill effects through early warning measures.

This paper deals with rainfall runoff modeling of the upper Awash River catchment located in the central part of Ethiopia. The area of the basin is 7615 sq. km. The river is the main tributary of the Koka multi purpose reservoir used for hydropower generation, irrigation and flood control.

The Galway Flow Forecasting System, GFFS, was used to model the flow. Various models of GFFS, ranging from simple linear models to SMAR – the conceptual model were used to simulate river flows. Brief description of the models is presented in the paper. Performance comparisons have been made to select models that are more suitable for the catchmenent. It was concluded that in the simulated mode, the SMAR model performs better than the others.

1. INTRODUCTION

The Awash Basin rises at an elevation of 3000masl over the central highland of Ethiopia about 150 km west of Addis Ababa and has a drainage area of 110,000 sq. kms. The river flows generally north eastwards along the Rift Valley and terminates in Lake Abe at an elevation of 250masl near the Djibouti border. The Koka reservoir is located in the upper reaches of the Awash basin approximately 75 km southeast of Addis Ababa, and has been in operation for the last 45 years. Commissioned in 1960, it was the first major hydroelectric development in the country. The 24 m high dam originally impounded some 1750 MCM of water. The dam provides a gross head of 42 m for the power plant, which has three units with total installed capacity of 43 MW. The Awash II and III plants lie in-series, down-stream from Koka, both have installed capacities of 32 MW each, with rated heads of 59.8m. Although, the dam was initially constructed for hydropower production, eventually it

becomes a multi purpose reservoir and serves as a source of water for the 30,000 ha irrigation farms down stream and for flood control.

In the present study the flow-gauging station of Awash River at Hombole was selected for the application and selection of suitable rainfall-rainfall models to forecast inflows into the reservoir. The drainage area of the station is 7615 sq. kms and the mean annual flow is about 1350 mcm, which is 90% of the total inflow into the reservoir and covers 70% of the total reservoir drainage area.

2. DATA AVAILABILITY

2.1 Climate

The mean annual rainfall of the basin is 1200 mm and reaches 1500 mm at the eastern highlands of Addis Ababa. The catchment receives its maximum rainfall during June-September and constitutes 70 to 75 % of the annual rainfall. The second rainy period covers the period from February to May. To minimize the error introduced in estimation aerial rainfall, due to high spatial variation, daily records for a total of nine-rainfall station located with in the catchment has been vollected. Thiesson Polygon method is applied to estimate the aerial rainfall. The location of these stations and the weighted areas are shown in fig 1.
Monthly Temperature data for five stations and Evaporation data for one station in the catchment have been used for the present modeling.

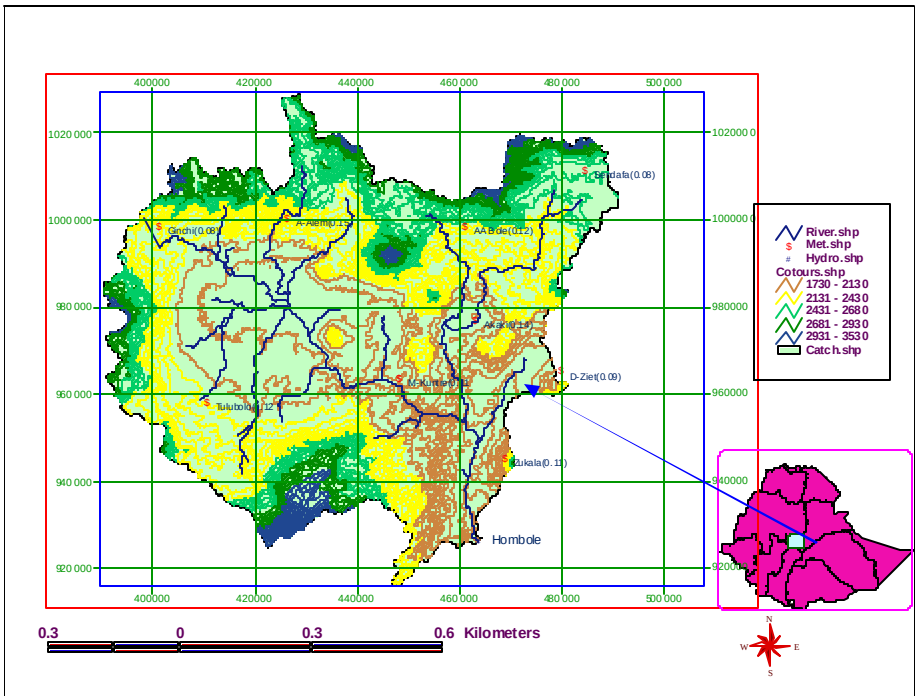


Fig. 1. Awash River at Hombole drainage area

2.2 Water Resources

In the study area there are a number of hydrological stations having daily records of ranging from few years to over 40 years. For the present study, a station located most down-stream on the Awash River above Koka reservoir is selected. The station has a drainage area of 7615 sq. kms. with mean annual flow of 1350 MCM. The flow of the river highly seasonal and 85% of the flow occurs during the months of June-September. The monthly rainfall, evaporation and flow depths are plotted and shown in fig. 2.

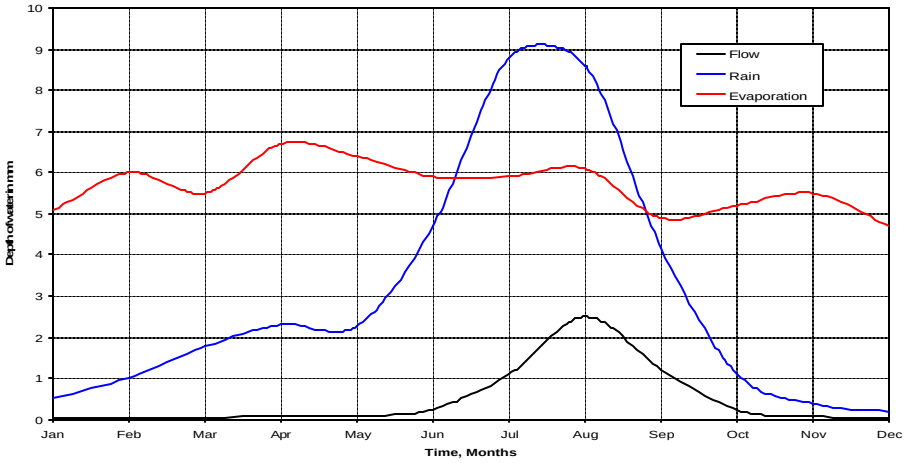


Fig.2 Monthly Rainfall, Evaporation and Flow of the Study area

2.3 Topography

Topographic maps at a scale of 1:50,000 are available for the area. Further, USGS Digital Elevation Model, DEM, of resolution approximately 90mX90m is available. The drainage area of the flow gauging station is delineated and basic drainage characteristics are also determined using the WMS software. The drainage area of the basin is 7615 sq. kms, the average elevation of the catchment is 2300 masl and ranges from 3500m to 1700 masl.

3. DESCRIPTION OF THE MODELS USED

For the present study the Galway Flow Forecasting System, GFFS, is used to model the catchment. The GFFS is a River flow forecasting system incorporating different hydrologic models. The Package is developed at the Department of Engineering Hydrology, National University of Ireland; Galway, Ireland. The models used in the GFFS are described below.

3.1 Simple Linear Models (SLM)

The intrinsic hypothesis of the SLM is the assumption of a linear time invariant relationship between the total rainfall R_i and the total discharge Q_i. In its discrete non-parametric form the SLM, including the forecast error term e_i, is expressed by the convolution summation relation [Kachroo and Liang, 1992],

$$Q_i = \sum_{j=1}^{m} R_{i-j+1} h'_j + e_i = G \sum R_{i-j+1} B_j \quad \text{where,} \quad \sum_{j=1}^{m} B_j = 1 \qquad \text{(EQ. 1)}$$

h'_j is the j-*th* discrete pulse response ordinate or weight, m is the memory length of the system, G is the gain factor and B_j is the arithmetic sum of the discrete pulse response ordinates which defines the gain factor G.

3.2 Linear perturbation Models (LPM)

In the LPM, [Nash and Barsi, 1983], it is assumed that, during a year in which the rainfall is identical to its seasonal expectation, the corresponding discharge hydrograph is also identical to its seasonal expectation. However, in all other years, when the rainfall and the discharge values depart from their respective seasonal expectations, these departures series are assumed to be related by a linear time variant system. Hence, the LPM structure reduces reliance on the linearity assumption of the SLM and gives substantial weight to the observed seasonal behavior of the catchment.

The relation between the departures/perturbation series of the LPM, incorporating an output error term ε_i, may be represented algebraically by the convolution summation equation of the same form as EQ 1 with Q_i and R_i are the respective departures of rainfall and discharge from their seasonal expectations. Model estimated departures values are added to the seasonal expectations to give the estimated discharge values.

3.3 Linearly Varying Gain Factor Model, LVGFM

The Linearly-Varying Gain Factor Model (LVGFM), proposed by [Ahsan and O'Connor, 1994] for the single input to single-output case, involves variation of the gain factor with the selected index of the prevailing catchment wetness, without varying the shape (i.e. the weights) of the response function. The model output has the familiar convolution summation structure (based the concept of a time-varying gain factor G_i

$$Q_i = G_j \sum_{j=1}^{m} R_{i-j+1} B_j, \; where \sum_{j+1}^{m} B_j = 1 \qquad \text{(EQ. 2)}$$

The multiples-input to single-output form of this model, was investigated by Liang et al. (1994). In its simplest form, G_i is linearly related to an index of the soil moisture state z_i of the catchment by the equation $G_i = a + bz_i$ (a and b being constants).

Although the antecedent precipitation index (API) provides a crude index of the current soil moisture state z_i ,[Ahsan and O'Connor 1994}, suggested that z_i be conveniently obtained from the outputs of the naive SLM, operating as an auxiliary modal, according to the relation,

$$z_i = \frac{\hat{G}}{Q} \sum_{j=1}^{m} R_{i-j+1} \hat{h}_j \qquad \text{(EQ. 3)}$$

Where, $\hat{G}$ and $\hat{h}$ are estimates of the gain factor and the pulse response ordinates respectively of the SLM and $\overline{Q}$ is the mean calibration discharge. The schematic diagram of LVGFM is shown in fig 3a.

3.4 Artificial Neural Network, ANN

The type of neural network used in the GFFS is the "multi-layer feed-forward network" which is considered to be very powerful in function modeling. It consist of an input layer, an output layer and only one "hidden" layer located between the input and the output layers, [Shamseldin, 1997]. A layer consists of a set of neurons each having the same pattern of connection pathways to the other neurons of the adjacent layers. Each neuron of a particular layer has connection pathways to all the neurons in the following adjacent layer, but none of those of its own layer or to those of the previous layer. Likewise, nodes in non-adjacent

layers are unconnected. Only one hidden layer is used in the version of the ANN included in the GFFS. The number of neurons in the input layer equals the number of the elements in the external input array to the network. There is only one neuron for the single output, in the out put layer.

As the neural network itself does not incorporate storage effects, storage is implicitly accounted for by the use of such an index. Instead of using the classic Antecedent Precipitation Index (API), involving a geometric weighting series, as the rainfall index, Shamseldin used the output series of the naive SLM. For a neuron either in the hidden or in the output layer, the received inputs Y_i are transformed to its output Y_{out} by a mathematical transfer function of the form:

$$Y_{out} = f\left(\sum_{i=1}^{M} w_i Y_i + w_0\right) \qquad \text{EQ. 4}$$

where, $f(\ )$ denotes the transfer function, w_1 is the input connection pathway weight, M is the total number of inputs (which usually equals the number of neurons in the proceeding layer), and W_0 is the neuron threshold (or bias). In the GFFS, the non-linear transfer function adopted for the neurons of the hidden layer and also that of the output layer is the widely used logistic function, i.e. a form of sigmoid function, given by:

$$F\left(\sum_{i=1}^{M} W_i Y_i + W_0\right) = \frac{1}{1 + e^{\,s\left(\sum_{i=1}^{M} W_i Y_i + W_0\right)}} \qquad \text{EQ. 5}$$

which is bounded in the range $[0,1]$ implying that the network output is likewise bounded in that range, s being a scaling parameter of the transfer function. The weights W_1, the threshold W_0 and s of the different neurons can all be interpreted as a parameters of the selected network configuration. The schematic diagram of ANN is shown in fig 3b.

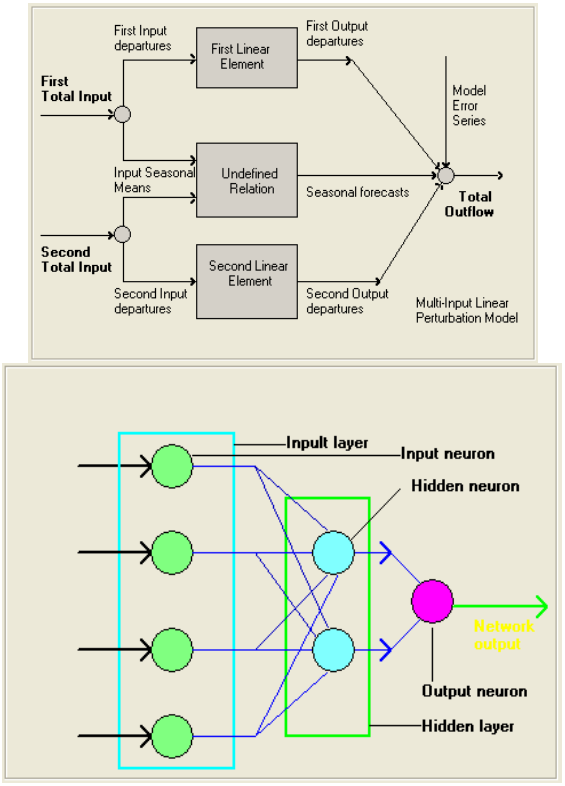
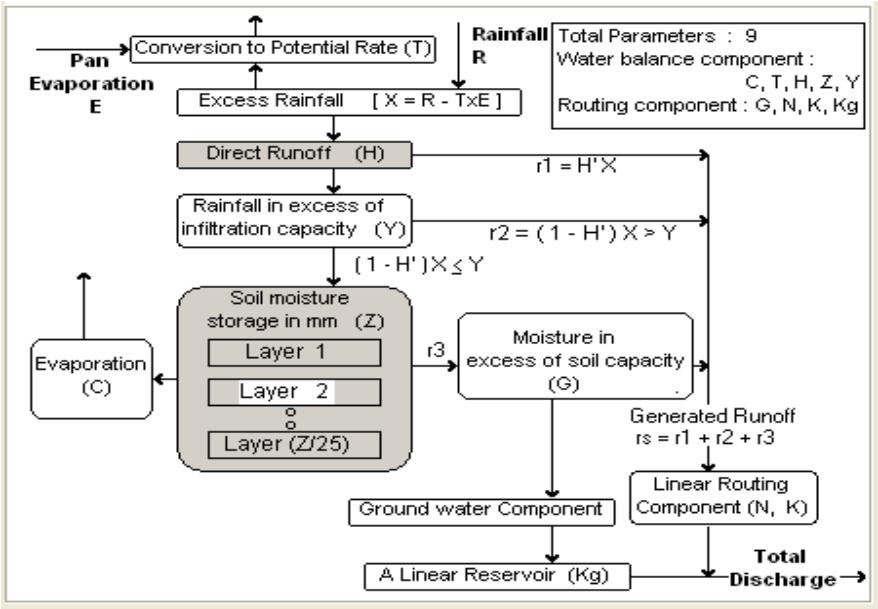


Fig. 3 Schematic Diagram of Linearly Varying Gain Factor (a) and Artificial Neural Network (b) Models

3.5 Soil Moisture Accounting and Routing, SMAR

The SMAR model is a lumped quasi-physical conceptual rainfall-evaporation-runoff model, with quite distinct water-balance and routing components. Using a number of empirical and assumed relations which are considered to be at least physically plausible, the non-linear water balance, i.e., soil moisture accounting, component ensures satisfaction of the continuity equation, over each time step. The routing component on the other hand, simulates the attenuation and diffusive effects of the catchment by routing the various generated runoff components, through linear time-variant storage elements. For each time-step, the combined outputs of the two routing elements adopted i.e., the generated surface runoff and groundwater runoff becomes the simulated discharge.

The version of SMAR used in this study has nine parameters, five of which control the overall operation of the water-budget component, while the remaining four parameters, including a weighing parameter, which determines the amount of generated groundwater runoff, control the operation of the routing component. The schematic diagram of the version of SMAR used in the present study incorporates the suggested modification of both [Khan, 1986 and Liang 1992]is the SMAR model along with the parameters used is presented in fig. 4.



Sl.no	Brief parameter description
1	T [Potential evaporation conversion coefficient]
2	H [Direct runoff separation coefficient]
3	Y [Soil moisture infiltration rate(mm/timestep)]
4	Z [Soil moisture storage capacity(mm)]
5	C [Evaporation decay coefficient]
6	G [Groundwater separation coefficient]
7	N [Linear reservoir nos in cascade]
8	NK [Time lag parameter for Nash cascade routing]
9	Kg [Time lag parameter for groundwater storage]

Fig. 4 Schematic Diagram of Soil Moisture Accounting and Routing Model

3.6 Models output Combination Techniques, MOCT

In MOCT, the outputs of the models described above are used as inputs to this model. The inputs are combined to give a single output using three combination techniques namely, Simple Average Method (SAM), Weighted Average Method (WAM), by applying different weights for each inputs and Neural Network Method (NAM).

3.7 Calibration and Verification in Updating Mode

The following model output updating techniques are used are: Auto Regressive (AR), Linear Transfer Function (LTF) methods and Non Linear Auto Regressive Exogenous input model using Neural Network (NARXM).

In the AR, time series model, the model is separately calibrated off-line to the error time series of the simulation mode discharge forecasts and subsequently it is used in real time for forecasting the errors in the simulation mode discharge forecasts, [Serban and Askew, 1991; Ashan and O'Connor, 2001}. These error forecasts are then simply added to the simulation mode discharge values to give the updated discharge.

The Auto regressive Exogenous-input (ARXM) also Known as Linear Transfer Function (LTF) model, [Shamseldin and O'Connor, 1999], is a linear input-output model which

enables the forecasting of the future values on the basis of the values of one or more exogenous input time series. In the updating procedure, the simulation mode discharge time series produced by the substantive rainfall-runoff model constitutes the exogenous input, which is used with the observed discharge in providing the updated discharge forecast.
In the NARXM, the neural network technique is applied for updating the discharge forecasts of the substantive rainfall-runoff model.

4. CALIBRATION AND VERIFICATION OF MODELS

All models in the GFFS described above were calibrated and verified for Awash River flow at Hombole station. Daily rainfall and flow records for the period starting Jan 1, 1991 up to Dec. 31, 2002 have been used. The data has been split in two periods for calibration and verification; from 1/1/1991 to 31/12/1998 for calibration and from 1/1/1999 to 31/12/2002 for validation.

The performances of the models are evaluated by visual comparison of the simulated and observed flows plots and by numerical efficiency criteria. The efficiency criteria used are:

- ***The Nash-Sutcliffe Criterion (1970) (R^2)***, which is related to the sum of the squares of the differences, F, between the estimated and observed discharges. The Criterion is defined by:

$$R^2 = \frac{F_0 - F}{F_0} \qquad F = \sum_{i=1}^{i=n} (Q_i - \hat{Q}_i)^2 \qquad F_0 = \sum_{i=1}^{i=n} (Q_i - \bar{Q}_i)^2 \qquad \text{EQ. 6}$$

where, Q_i is Observed Discharge, $\hat{Q}_i$ is Estimated Discharge, $\bar{Q}$ = Mean of observed discharge and F_o and F are the initial and final variance respectively.
The value of R^2 normally varies between 0 and 1 with higher values indicating better agreement.

- ***Index of Volumetric Fit (IVF),*** is the ratio of the total estimated volume Q_e, to the total observed volume Q_o, and is expressed as:

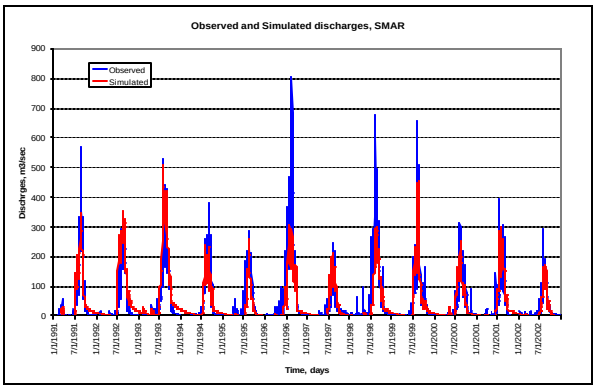
$$IVF = \sum_{i-1}^{N} (Q_e)_i \,/\, \sum_{i-1}^{N} (Q_o)_i \qquad \text{EQ. 7}$$

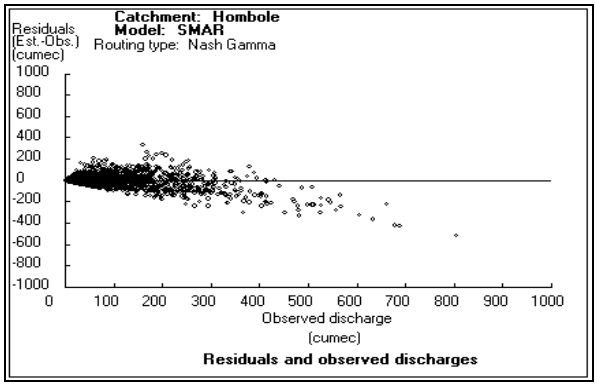
- ***The Relative error to the peak*** **(RE),** is defined as

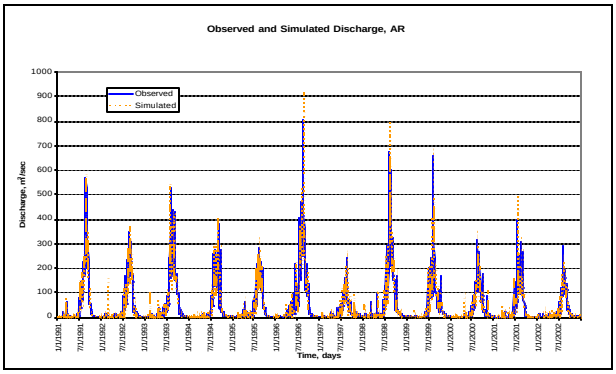
$$RE = \frac{\left|(Q_P)_e - (Q_P)_o\right|}{(Q_P)_o} \qquad \text{EQ. 8}$$

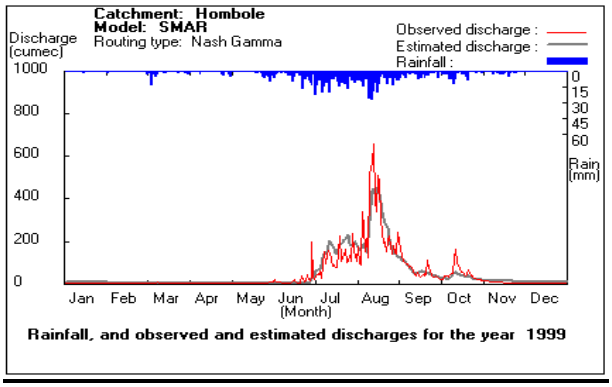
where, $(Q_P)_o$ and $(Q_P)_e$ being the observed and estimated peak flows respectively

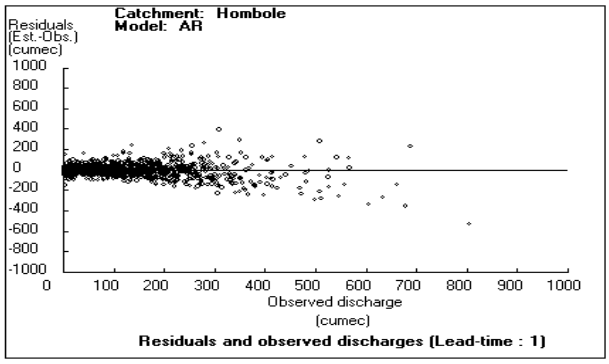
The observed and simulated discharges for the whole period has been plotted and shown from fig. 5. The model efficiencies obtained for different model are presented in tables 1 and 2.


Observed and Simulated discharges, SMAR
Observed
Simulated
Discharges, m3/sec
Time, days


Catchment: Hombole
Model: SMAR
Routing type: Nash Gamma
Residuals (Est.-Obs.) (cumec)
Observed discharge (cumec)
Residuals and observed discharges


Observed and Simulated Discharge, AR
Observed
Simulated
Discharge, m³/sec
Time, days





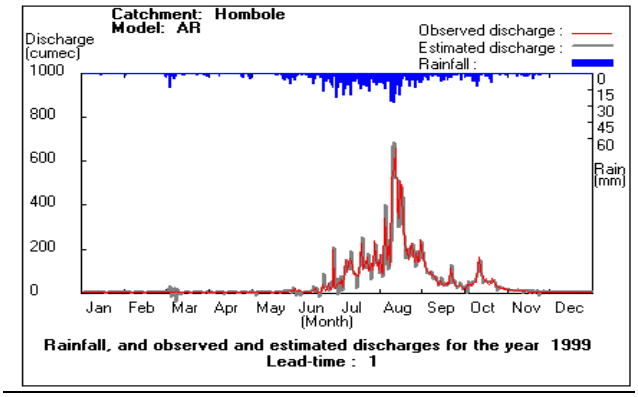


Figure 5. Plots of Observed and Simulated discharges and residuals of selected models

Table 1 Model Efficiencies in Simulation Mode

Model	Calibration			Verification			Remark
	R^2	IVF	RE	R^2	IVF	RE	
NPSLM	0.52	1.10	0.67	0.47	1.40	0.47	
PSLM	0.48	1.04	0.62	0.4	1.34	0.44	
NPLPM	0.72	0.99	0.65	0.64	1.32	0.52	
PLPM	0.76	1.00	0.56	0.59	1.33	0.45	
LVGFM	0.53	1.01	0.56	0.55	1.25	0.20	NPSLM
SMAR	0.72	1.02	0.37	0.79	1.09	0.37	
ANN	0.51	0.98	0.74	0.39	1.24	0.68	
MOCT_SAM	0.74	1.01	0.52	0.78	1.22	0.35	The inputs
MOCT_WAM	0.77	1.02	0.5	0.77	1.23	0.43	are outputs
MOCT_NNM	0.80	1.00	0.53	0.79	1.19	0.44	of NPLPM, SMAR and LVGFM models

Table 2. Model Efficiencies in Updating and Mode

Model	Calibration			Validation			Remark
	R^2	IVF	RE	R^2	IVF	RE	
AR (LT 1)	0.87	1.00	0.15	0.87	1.01	0.04	LVGFM,
AR (LT 3)	0.69	1.00	0.04	0.69	1.03	0.24	2^{nd} order
LTF (LT 1)	0.87	0.99	0.21	0.85	1.06	0.11	SMAR
LTF (LT 3)	0.81	1.05	0.38	0.79	1.14	0.17	
NNU	0.77	1.21	0.16	0.78	1.24	0.16	SMAR
NARXM	0.86	1.05	0.56	0.81	1.07	0.47	

Note: R^2 = Coefficient of efficiency (Nash and Sutcliffe, 1970)
IVF = Index of Volumetric Fit
RE = Relative Error of the peak flows
LT1 and LT3 are lead-time of 1 and 3 days respectively.

5. DISCUSSION ON THE RESULTS

The Simple Linear Model, both the parametric and non-parametric, results are inferior to that of other models. The R^2 values are low and they could not estimate the peak flows. The LPM, with its inherent component of seasonal variation, performs well in terms of R^2 and IVF during calibration period. But these performances deteriorate in the verification period. The ANN, although characterized by large number of weights does not perform well. The SMAR model, being conceptual and takes into account the soil moisture condition, performs better than all the models in all performance criteria used in both simulation as well as verification mode.

Although, strictly the MOCT is not a rainfall-runoff model, the combination techniques incorporated in it, especially the NNM techniques give better results.

In the model output-updating procedures, the AR updating procedure performs best than the others in updating the model outputs specifically for the lead-time of one day. The performance falls sharply as the lead-time increases. The LTF followed by the NARXM gives very good results.

For the objective the modeling work, i.e., to forecast inflows, volumes and peak flows, to the Koka reservoir for better reservoir management, the SMAR model coupled with AR updating procedure is suitable to forecast reservoir inflows (volumes) for reservoir management and peak flows for the operation of the spillway gates to avoid huge releases through the gates which might causes floods down-stream.

6. RECOMMENDATION

Since the catchment is heterogeneous land use/cover characteristics, considering the whole catchment as one unit could affect the performances of the SMAR lumped conceptual model. In order to avoid errors resulted in lumping the catchment characteristics, it is recommended divide the catchment to sub-catchments and to use distributed conceptual models.

7. Acknowledgement

This paper was prepared based on the research activities of the FRIEND/Nile Project which is funded by the Flemish Government of Belgium through the Flanders-UNESCO Science Trust Fund cooperation and executed by UNESCO Cairo Office. The authors would like to express their great appreciation to the Flemish Government of Belgium, the Flemish experts and universities for their financial and technical support to the project. The authors are indebted to UNESCO Cairo Office, the FRIEND/Nile Project management team, overall coordinator, thematic coordinators, themes researchers and the implementing institutes in the Nile countries for the successful execution and smooth implementation of the project. Thanks are also due to UNESCO Offices in Nairobi, Dar Es Salaam and Addis Ababa for their efforts to facilitate the implementation of the FRIEND/Nile activities.

8. References

Ashan, M. and O'Connor, K.M.. 1994b. A simple non-linear rainfall-runoff model with a variable gain factor. J.Hydrol., 155: 151-183.

David R. Legates and Gregory J. McCabe Jr., 1999. Evaluating the use of "goodness-of-fit measures in hydrologic and hydroclimatic model validation. Water Resources Research,Vol. 35, No1, 233-241

Goswami, M., O'Connor, K.M. and Shamseldin, A.Y., Structures and Performances of Five Rainfall-Runoff Models for River-Flow Simulation. 476-480

Kachroo, R.K. and Liang, G.C., 1992. River Flow Forecasting. Part 2. Algebraic development of linear modeling techniques. J.Hydrol., 133: 17-40

Khan, H., 1986. Conceptual Modeling of rainfall-runoff systems. M.Eng. Thesis, Nat. Univ. of Ireland, Galway.

Liang, G.C., 1992. A note on the revised SMAR model. Workshop Memorandum, Dept. of Engineering Hydrology, Nat. Univ. College of Ireland, Galway, (Unpublished).
Nash, J.E. and Barsi, B.I.,1983. A hybrid model for flow forecasting on large catchments. J. Hydrol., 65: 125-137

Nash, j.e. and Sutcliffe, J.V.,1970. River flow forecasting through conceptual models. Part 1. a discussion of priciples. J.Hydrol., 10, 282-290.

Serban, P, and Askew, A.J., 1991. Hydrological forecasting and updating procedure. IAHS pub. No. 201, 357-369.

Shamseldin.A.Y.,1997. Application of a neural network technique to rainfall-runoff modeling, J.Hydrol., 199, 272-294.

Shamseldin A.Y. and O'Connor, K.M.., 2001. A non-linear neural network technique for updating of river flow forecasts. Hydrology & Earth system Sciences, 5(4), 577-597