Development of the SPLASH model:
A deterministic conceptual daily rainfall-runoff model for Southern Africa

S. van Biljon
SADC-HYCOS Project Management Unit, PO Box 32798, Glenstantia, Pretoria, 0010, South Africa
Tel: +2712 841 3219; Fax: +2712 841 3228; e-mail: svb@sadchycos.co.za

1. Introduction

The development of surface water resources for the benefit of man requires a thorough study of streamflow time series. Since the measurement of runoff in rivers is difficult and expensive, the availability of streamflow time series at the sites of interest, even in modern times, leaves much to be desired. Originally only historical streamflow traces were studied, but with the increasing costs associated with the construction of water schemes and the inevitable demands and effects of increasing populations, complexities of analytical hydrological techniques soon escalated resulting in modelling techniques both in the deterministic, statistical and stochastic domain. This paper will discuss deterministic topics with an example of a conceptual daily rainfall-runoff model.

2. Definitions and classifications

For the purposes of this paper, it is necessary to define terminology often used in hydrological modelling. A *variable* is defined as a measurable entity and takes on a *different value at different times*. If this entity is measured at fixed time intervals it constitutes a *time series*. Daily and monthly rainfall and runoff values are good examples of hydrological time series. The hydrological cycle is known to consist of complex physical, chemical and biological processes. A *model* may be regarded as a *simplified representation of the hydrological cycle*, mostly describing only part of the cycle where certain variables are used as input to obtain output variable(s). *Parameters* are defined as *time invariant* quantities to convert input in models to the required output. Chow (1964) said that if the *chance of occurrence* of variables is *ignored* and a model is considered to follow a definite law of certainty, but not any law of probability, the *process and its model* are described as *deterministic*. Literature abounds with model type definitions. The most common of these are the black box or empirical models (also referred to as input-output models) and conceptual models. Some models are further typified as linear or non-linear. When the input variables to a model do not take account of its spatial distribution such a model is called a lumped model. A distributed model on the other hand describes spatial variability.

Linsley (1981) classified rainfall-runoff models as deterministic, stochastic, conceptual, theoretical, black-box, continuous, event, complete, routing or simplified. Existing models can be categorized according to Linsley's classification. It will also become clear that specific models may be classified according to more than one attribute.

Song and James (1992) identified five different scales at which hydrological modelling may take place, *i.e.* laboratory-scale, hillslope-, catchment-, basin- and continental or global-scale. Essentially the *hillslope*-scale simulates hydrological behaviour on a very small catchment and only differs from the *laboratory* scale in that certain physical properties such as infiltration, sub-surface flow, soil saturation and the transfer to overland flow cannot be replicated accurately in the laboratory. Models for operational purposes such as the Stanford Watershed Model (Crawford and Linsley 1966) and Pitman (1972) model operate on the *catchment* scale where for instance geology, topography and time of concentration plays a role. Models at the *basin*-scale add complexity to catchment-scale models since peak flow attenuation takes place through storage and routing along rivers. *Continental* or *global*-scale models primarily focus on atmospheric processes that govern evapotranspiration and precipitation.

3. Deterministic modelling

Generally, the smaller the scale for modelling is, the more intense the data requirements will be for modelling a particular subset of the hydrological cycle. Even at catchment scale, models such as ACRU (Shulze 1989), which was originally developed for very small catchments, and the Pitman model (1973, 1991) have distinctly different data requirements. The former model can be described as a non parametric distributed physical model which simulates hydrological behaviour of the rainfall-runoff process from extensive data on soils, evaporation, rainfall, topography, land use *etc*. By aggregating the modelling results from very small catchment units, runoff for larger catchments are obtained. Data assembly for the ACRU model is time consuming and best handled by means of a Geographical Information System (GIS). The Pitman model has been developed from concepts of the Stanford Watershed Model as described by Crawford and Linsley (1966) and can be described as a lumped conceptual rainfall-runoff model. The structure, calibration, verification and application of these models have been extensively documented in the literature and will not be discussed any further.

Certain practical problems haunt successful deterministic modelling. One of the most important facts concerns the spatial rainfall input to these models. On the catchment and basin scale point rainfall measurement has to be lumped or extrapolated on a gridded basis to represent spatial rainfall (Lynch 1991, Seed 1992). Since very few autographic rainfall measurements are available, extrapolations are being made from point rainfall data to obtain catchment rainfall. The true temporal and spatial rainfall picture is largely unknown. In mountainous catchments where the highest rainfall and runoff occur, neither is adequately monitored. Rainfall shadows are largely unknown and have to be determined from models and practical knowledge of the area. No model, regardless how accurate in theory or concept it may be, can yield better performance than the input variables allow.

Catchment management practices present another very significant problem in modelling. The influence of irrigation, urban areas, return flows, afforested areas, agricultural developments and especially the construction of small farm dams are dynamic variables which perturb the natural behaviour of the runoff process. For this reason it is very important to study the catchment and its management over time and to reconstruct the runoff time series to account for these effects before models can be successfully calibrated or fitted. Data on these anthropogenic aspects are unfortunately lacking for most of the catchments and use has to be made of growth functions to account for these whenever historical data becomes available from whatever source. The addition (or subtraction) of constant seasonal or annual flow values to (or from) the observed runoff series often introduce more errors into the time series and should be discouraged. Factors such as the historical rainfall regime, economic conditions, state of vegetation and agricultural production, all form complex relationships which ideally should be studied to obtain a stochastic historical consumptive-use record. It is best to add such a record to the observed runoff series for modelling the catchment in its derived pristine condition. Present day or future levels of runoff perturbation should be reintroduced in the runoff series before further analyses for systems yield, stochastic modelling or whatever purpose runoff traces are needed.

4. Description of a lumped conceptual daily rainfall-runoff model

Large reservoir systems are usually analyzed on a monthly time scale, primarily to accommodate seasonal runoff, evaporation and demand characteristics. Run-of-river diversion is best accounted for on a daily basis. The importance of the environment has reached a stage where certain phenomena have to be studied with at least a daily time resolution. Flood forecasting for large systems, (like the Vaal and Orange River systems in South Africa) also require at least daily information. For these reasons, a literature survey was conducted by the author to increase his understanding of different models and also to determine the availability of daily rainfall-runoff models. Tremendously complicated models are available world wide, many of which require an enormous amount of input data. Particular attention was paid to reasonably simple models in terms of structure, input and parsimony of parameters.

From experience it is known that significant runoff events are either produced from extraordinary big storms or from (smaller) rainfall events on saturated catchments. A model first described by Nielsen and Hansen (1973) and later slightly revised by Refsgaard and Havnoe (1983) for flood prediction, provided the opportunity for further investigation. The author experimented with the original concepts which were developed for Danish hydrological conditions and found the results to be encouraging. A computer program was developed where certain modifications to the original models were introduced to yield acceptable results for South African hydrological conditions. It was tested using data from the Klaserie River at Fleur de Lys in the Mpumalanga Province of South Africa. The catchment is situated on the eastern slopes of the Drakensberg escarpment. The catchment area of the gauging site B7H004 is 136 km^2. A description of the model, called SPLASH (Simplified Precipitation Lumped Algorithms for Streamflow Hydrographs), is given below.

5. Description of the SPLASH model

The model concept of a catchment is a set of four different storage zones: the upper and lower zone, together with two groundwater zones. Evaporative losses from the upper zone are assumed to occur at the same rate as the potential evaporation. Once the moisture store in the upper zone is depleted, evaporative losses are assumed to take place from the lower zone according to a non-linear function of soil moisture. Overland flow and interflow are derived from the upper zone. When the lower zone storage is lower than a predetermined value, overland flow is determined by the storage level of the upper zone only. When the predetermined upper zone storage level is exceeded both overland and interflow are controlled by the moisture level in the lower zone. The two groundwater stores are replenished from the upper zone, the amount also controlled by the moisture store in the lower zone. The groundwater stores contribute to the baseflow.

The model is characterised by a mixture of linear and non-linear functions. Figure 1 gives a schematic representation of the model.

6. Structure of the SPLASH model and its parameters

*Upper zone storage U**

U* is the maximum storage (mm) of the upper zone and the active level of storage is U. U is replenished by rainfall and depleted by evaporation which takes place at a maximum rate equal to the potential evaporation. Potential evaporation is assumed to be equal to A pan evaporation and supplied as mean monthly values.

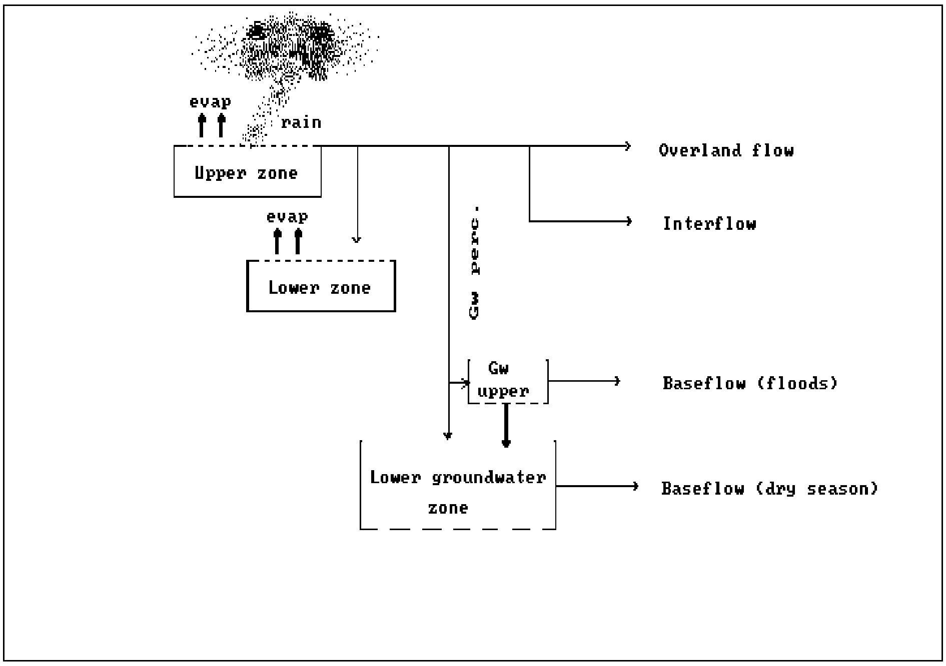


Figure 1: Schematic representation of the SPLASH model.

Overland flow (OF), interflow (IF), infiltration to the lower zone (DL) and percolation to the two groundwater zones (G) are derived from the upper zone in any combination or simultaneously on any given day provided U > U*. The occurrence of the latter processes are controlled by the moisture level in the lower zone.

Interflow is assumed to be proportional to U and to vary linearly with the relative moisture content L/L* of the lower zone. The relationship is given by the following formula (parameters will be described later in the document):

$$IF = CIF*(L/L^* - CLI)*U / (1 - CLI) \text{ for } L/L^* > CLI$$
$$= 0 \text{ for } L/L^* <= CLI$$

Overland flow is assumed to be proportional to the excess rainfall (PN) where PN = Rainfall + U - (evaporation + U*):

$$OF = COF*(L/L^* - CL2)*PN / (1 - CL2) \text{ for } L/L^* > CL2$$
$$= COF*(U/U^* - CL2)*PN / (1 – CL2) \text{ for } L/L^* <= CL2$$

Infiltration to the lower zone is given by DL and percolation to the groundwater zones by G. The respective relationships are the following:

$$G = (PN - OF)*(L/L^* - CLG) / (1 - CLG) \text{ for } L/L^* > CLG$$
$$= 0 \text{ for } L/L^* <= CLG$$

and $DL = (PN - OF) - G$

The smaller U* the greater DL will be for a given rainfall. This will give higher overland flows, higher baseflow and higher lower zone storages which will increase interflow. There is an offset, since the higher overland flow, the less the available water for infiltration will be.

The upper and lower groundwater zones contribute to flood and baseflow respectively. The relationships are the following:

$$BFU = (BFO*\exp(-1/BKU)) + G*(1 - CBL)*(1 - \exp(-1/BKU))$$
where BFO = BFU the previous day and

$$BFL = (BFLO*\exp(-1/BKL)) + G*CBL*(1 - \exp(-1/BKL))$$
where BFLO = BFL the previous day.

*Lower zone storage L**

The lower zone moisture level L is replenished by infiltration from the upper zone until the maximum storage level L* is reached. Excess moisture is routed to the groundwater stores. High rainfall will give rise to a number of such occurrences but a high frequency of such warnings may indicate that L* is set too small. Moisture from the lower zone is depleted by evaporative loss only.

The smaller L* the easier it will be to fill the lower zone storage with the result that the ratio L/L* will be higher. Since this ratio controls most of the runoff processes, interflow will be higher, overland flow will be higher and more water will percolate into groundwater storage with the result that baseflow will be higher.

Overland flow coefficient COF

COF determines the amount of overland flow should U > U*. This coefficient is further modified by CL2 conditional upon the active storages in the upper and lower zone.

Higher COF will increase overland flow and less water will be available to infiltrate, thus lower zone replenishment will be less and baseflow less.

Secondary overland flow coefficient CL2

The effect of CL2 is less pronounced than COF. However, CL2 performs an important task since it controls the level of storage in the lower zone at which overland flow will occur. The higher CL2 is set during calibration, the lower OF will be; or alternatively, the lower the frequency of overland flow occurrences will be. Lower OF will result in (PN – OF) being greater. (PN - OF) controls the infiltration and thus the contribution to the lower zone (which affects L/L*) and G which controls baseflow. Since L/L* controls also interflow, it is expected that interflow will increase.

Interflow coefficient CIF

Similar to COF, CIF controls the amount of interflow.

Higher CIF will increase interflow provided L/L* > CLI and U > 0. Depletion of upper zone storage U will thus be faster resulting in DL being lower. This will lower the contribution to infiltration affecting the lower zone (and thus L/L* which affects the whole process) and lower the baseflow.

Secondary interflow coefficient CLI

An increase in CLI will lower IF. Lower IF will deplete upper zone storage U slower; however U is in the first line of attack for fulfilling the demand for evatranspiration with the result that less interflow will occur and more rainfall input is required to replenish U, consequently P will be slightly lower with slightly lower G and thus baseflow.

CLI controls the level at which interflow occurs; the higher CLI, the less the frequency of interflow.

Time constants FKO and FKI

FKO and FKI are the time constants for overland and interflow respectively. It has virtually no effect on the amount of runoff but controls the time distribution (or lag) of overland and interflow.

Baseflow parameters CLG and CBL

CLG defines the threshold (in terms of the ratio L/L*) when deep percolation into the groundwater store commences. CBL defines the ratio of total groundwater percolating into the lower zone. Together with BKU and BKL these parameters have a significant effect on simulated results.

Baseflow time constants BKU and BKL

Splitting the groundwater storage into two zones resulted in better hydrograph shape definition. The baseflow is assumed to follow an exponential decay with replenishment from infiltration. BKU defines the upper groundwater zone and primarily defines the short term characteristics of flood hydrographs.

BKU and BKL are defined as the time constants of the groundwater reservoirs. BKL may be estimated from the recession of the streamflow hydrograph for the catchment under study with long periods of no rain, i.e. baseflow characteristics of the winter in the summer rainfall region or summer flows in the winter rainfall region.

These parameters have a significant influence on the model performance since the higher BK values, the longer baseflow persists with the effect that the total baseflow is higher.

Evaporative loss from the lower zone: CEL coefficient

The only depletion of moisture from the lower zone takes place by evapotranspiration losses. Due to long dry spells of virtually no rainfall during the winter months in the summer rainfall region, it was found that evaporative losses effectively emptied the lower zone moisture store. This required huge or prolonged rainfall input to replenish the lower zone storage with the model consistently under predicting runoff. Since it is physically impossible for evapotranspiration to continue unabatedly from underground, CEL controls evaporative losses through a non linear function. It is recommended to set CEL = 1 when calibration commences. Increasing CEL will effectively retard evaporative loss resulting in higher runoffs being simulated.

This is accomplished by calculating the evaporation loss from the lower zone as follows:

$$E_{lz} = E*(1/\exp(CEL))*\exp(CEL*L/L*)$$

where E is that part of evaporative loss from the lower zone to fulfil the required evapotranspiration loss by the atmosphere and E_{lz} the inhibited (or modified) loss.

7. Calibration and validation of the SPLASH model

Model development, calibration and validation ideally require catchments in pristine condition with reasonably long rainfall and runoff time series. The catchment selected for this purpose is the Klaserie River at Fleur de Lys, gauging station number B7H004, situated in the Drakensberg escarpment in Mpumalanga Province in South Africa. The catchment area is 136 km^2 and drains part of the eastern slopes of the Drakensberg Mountain range. The mean annual rainfall is approximately 1300 mm. Three rainfall stations of the South African Weather Service (SAWS) were used (with equal weighting) to define the catchment rainfall. Daily rainfall and runoff data for the hydrological years 1970/71 to 1979/80 were used to calibrate the model for the Klaserie River at Fleur de Lys.

Considering that the point rainfall measurements do not necessarily define the spatial rainfall of any catchment correctly (it should rather be seen as a catchment rainfall index), the overall fit of the SPLASH model is considered to be good to excellent. Figure 2 shows the simulated and measured daily runoff for the 1971/72 hydrological year. The largest observed flood peak in February 1972 was simulated very well.

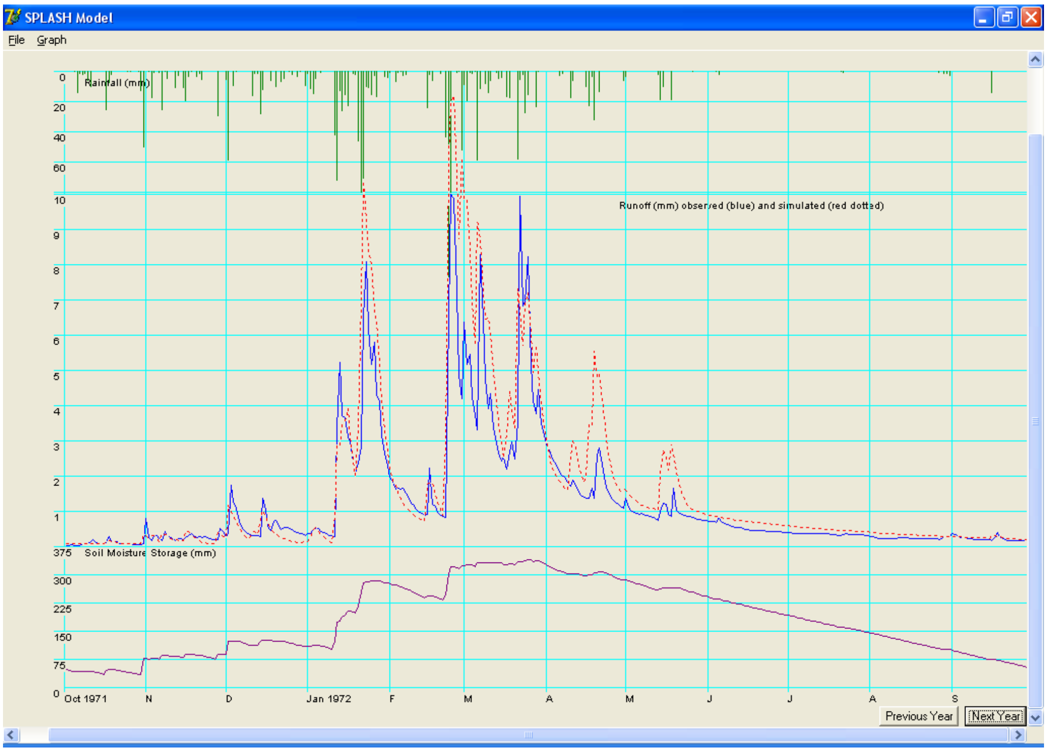


Figure 2: Model estimate of the flood peak at the end of February 1972 is higher than gauging weir / calibration table limit

Figures 3 and 4 display the application of the model to fill in gaps within the observed record. The largest February flood peaks simulated and shown in figures 3 and 4 are probably much better estimates than the peaks contained within the flow record which was measured either up to the gauging weir or calibration table limit.

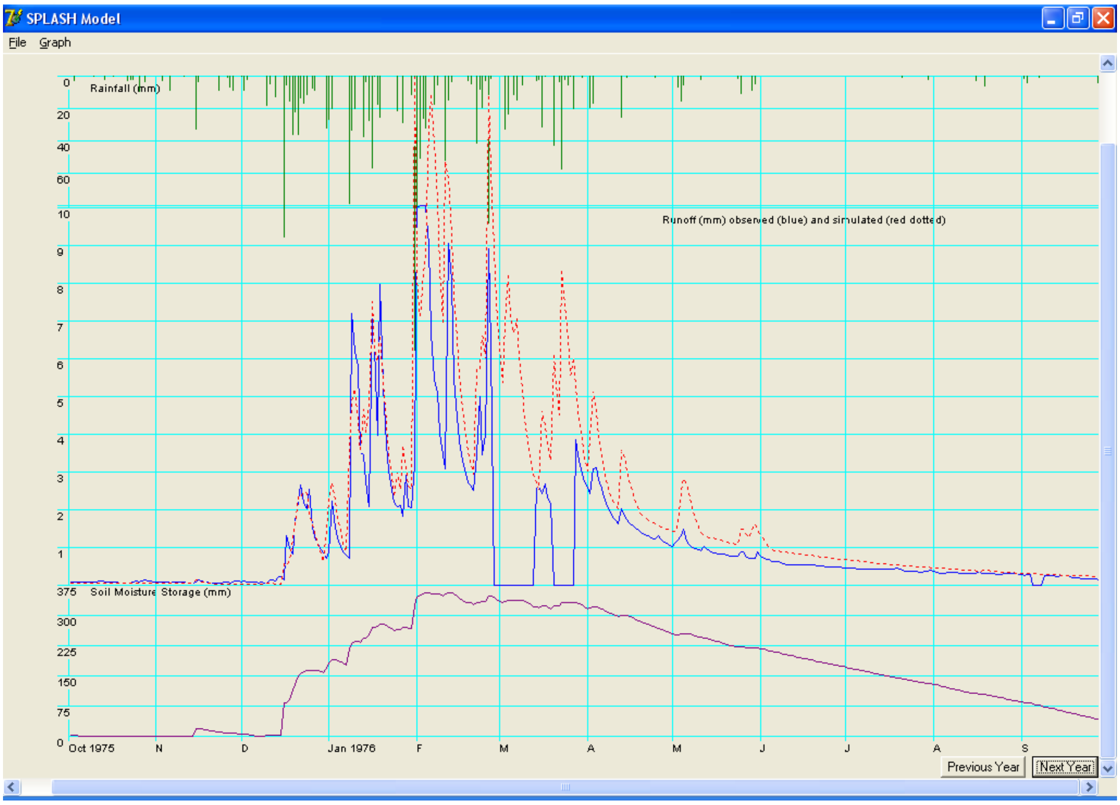


Figure 3: Weir / calibration table limit exceeded early in February 1976. Note the in-filled flows during Feb/Mar 1976

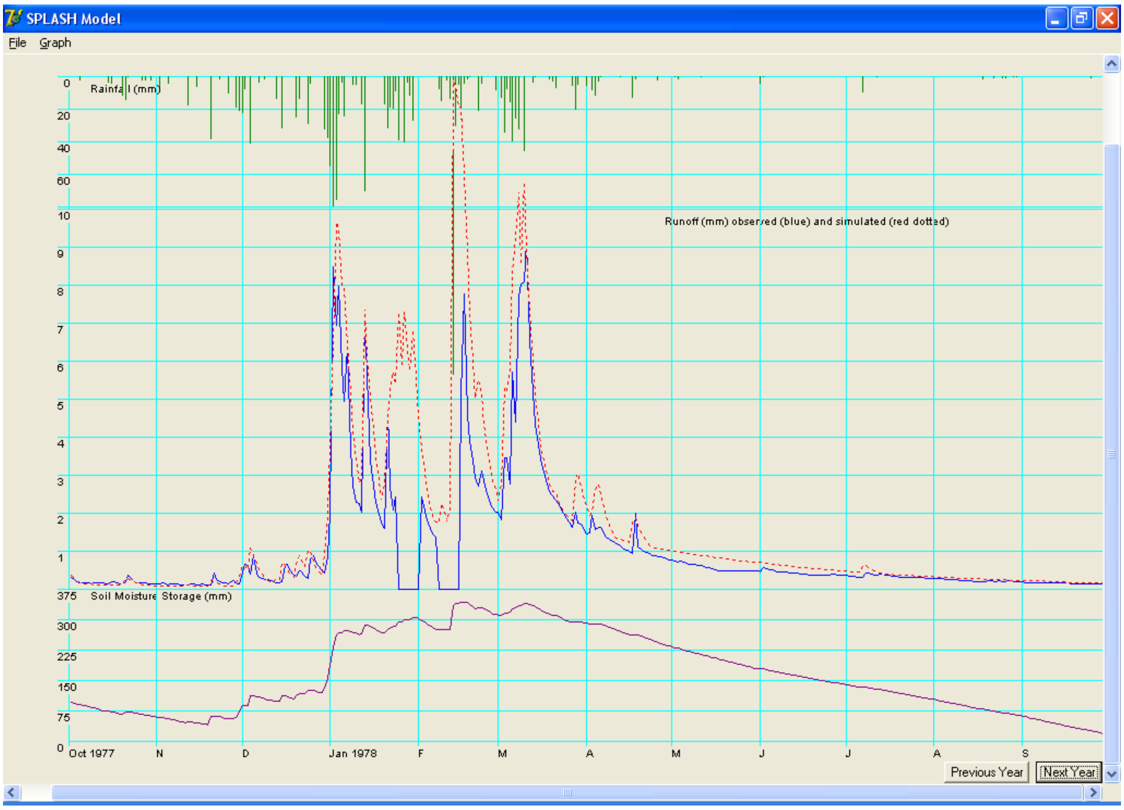


Figure 4: Note the SPLASH model estimates for periods with no record

Figure 5 shows the model performance for a dry year. The early part of the hydrological year is consistently under predicted. Once the lower zone storage has been partly replenished, modelled runoff appears to be more accurate. The correlation between the measured and simulated daily runoff was calculated for each of the 10 years and is given in Table 1. (No correlations were done during periods of missing observed records).

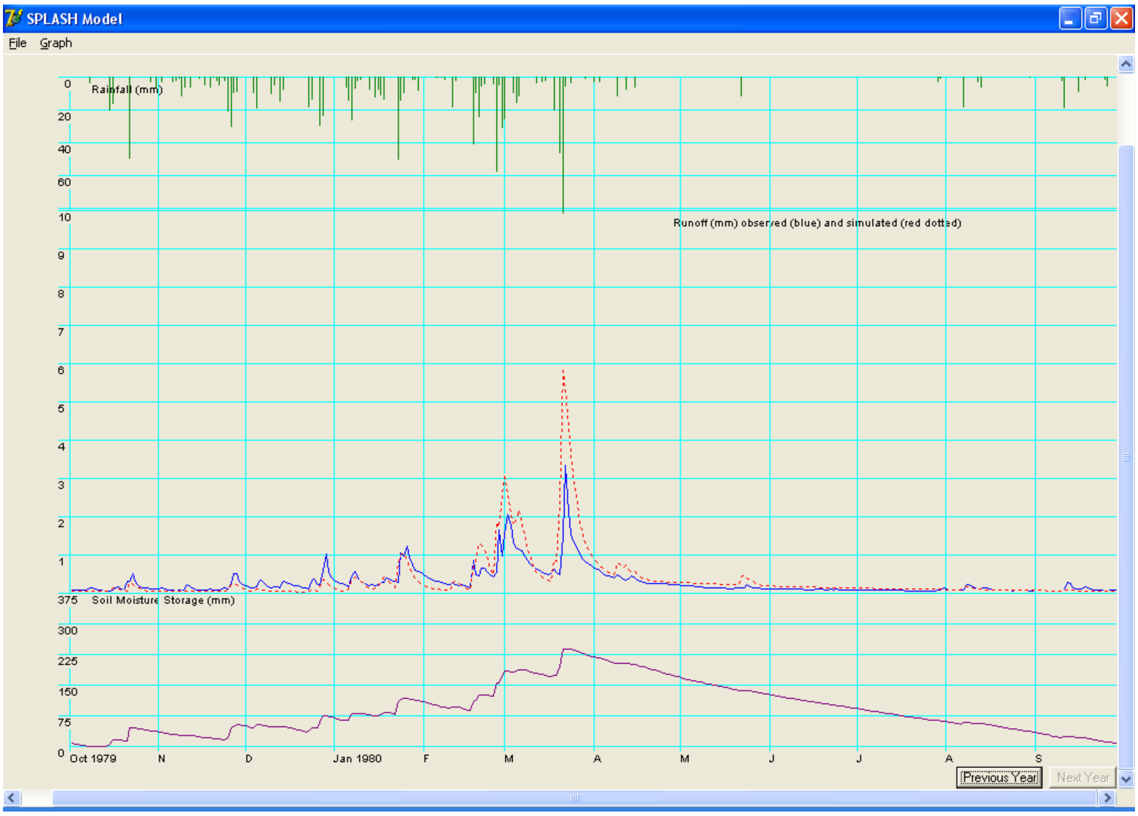


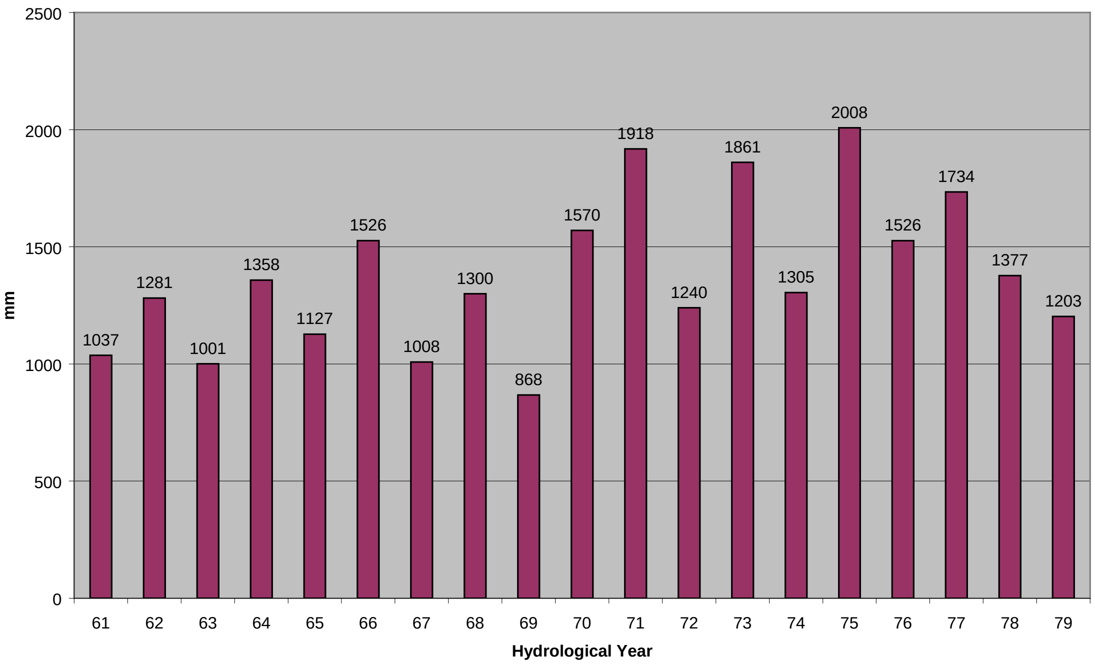
Figure 5: SPLASH model performance for year with late rains: initial under estimation; improvement with LZS replenishment.

Table 1: Comparative statistics for the Klaserie and Vaal Rivers: Shaded areas refer to the calibration period.

Klaserie River at Fleur de Lys No. B7H004 Catchment area: 136 km^2			Vaal River at Standerton No. C1H001 Catchment area: 8193 km^2		
Hydro Year	**Rainfall (mm)**	**r** $_{\text{obs vs sim}}$	**Hydro Year**	**Rainfall (mm)**	**r** $_{\text{obs vs sim}}$
1960/61			1960/61	712	0.62
1961/62	1037	0.67	1961/62	482	0.65
1962/63	1281	0.75	1962/63	545	0.87
1963/64	1001	0.55	1963/64	544	0.69
1964/65	1358	0.73	1964/65	587	0.81
1965/66	1127	0.54	1965/66	353	0.53
1966/67	1526	0.72	1966/67	764	0.40
1967/68	1008	0.63	1967/68	517	0.53
1968/69	1300	0.59	1968/69	621	0.36
1969/70	868	0.69	1969/70	588	0.66
1970/71	1570	0.75	1970/71	506	0.52
1971/72	1918	0.86	1971/72	663	0.66
1972/73	1240	0.88	1972/73	572	0.52
1973/74	1861	0.81	1973/74	664	0.60
1974/75	1305	0.85	1974/75	817	0.74
1975/76	2008	0.82	1975/76	739	0.63
1976/77	1526	0.73	1976/77		
1977/78	1734	0.88	1977/78		
1978/79	1377	0.77	1978/79		
1979/80	1203	0.79	1979/80		

For model validation the model parameters obtained during calibration were used for simulating the runoff for the period 1961/62 to 1969/70. Figure 6 illustrates that the daily rainfall for the validation period is significantly lower than the rainfall during the calibration period. Despite this appreciable difference in hydrological regime, simulated runoff was found to be acceptable. Correlation coefficients are given in Table 1. Figures 7 and 8 display the runoff generated for the wettest and driest year respectively.

FIGURE 6: Klaserie River Annual Catchment Rainfall



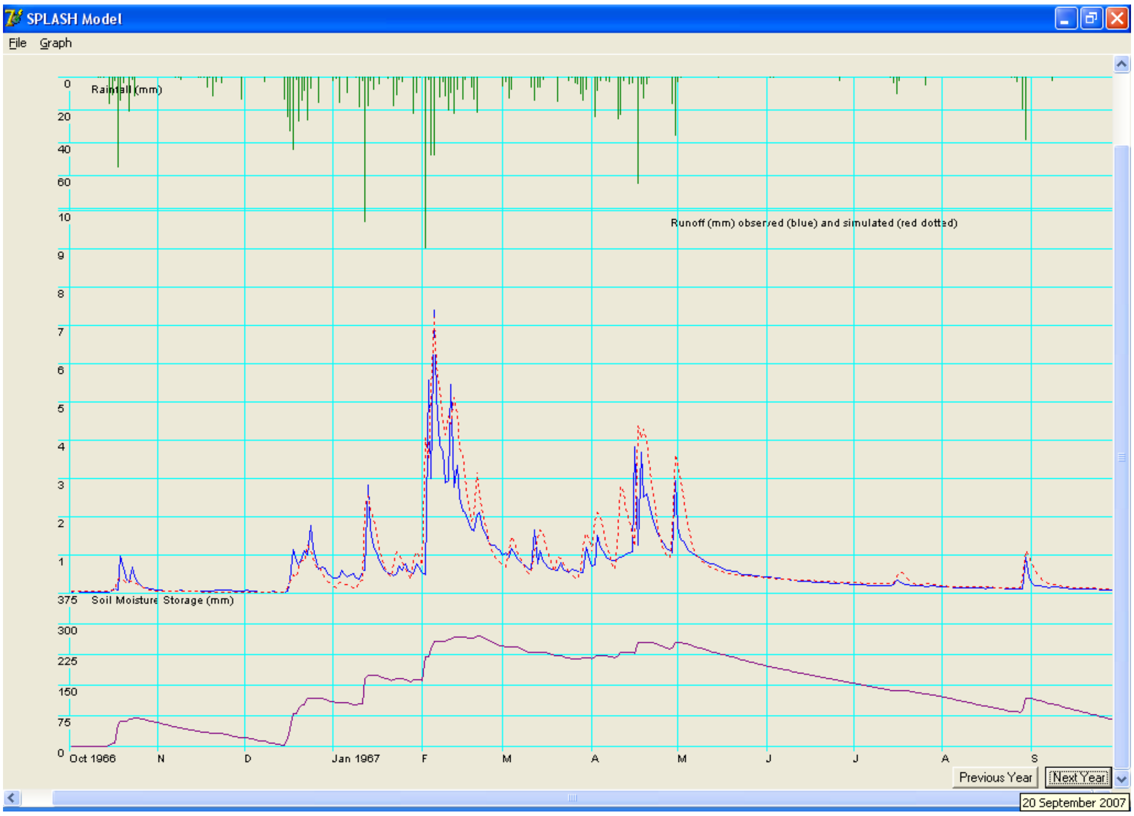


Figure 7: Model performance for the wettest year outside the calibration period.

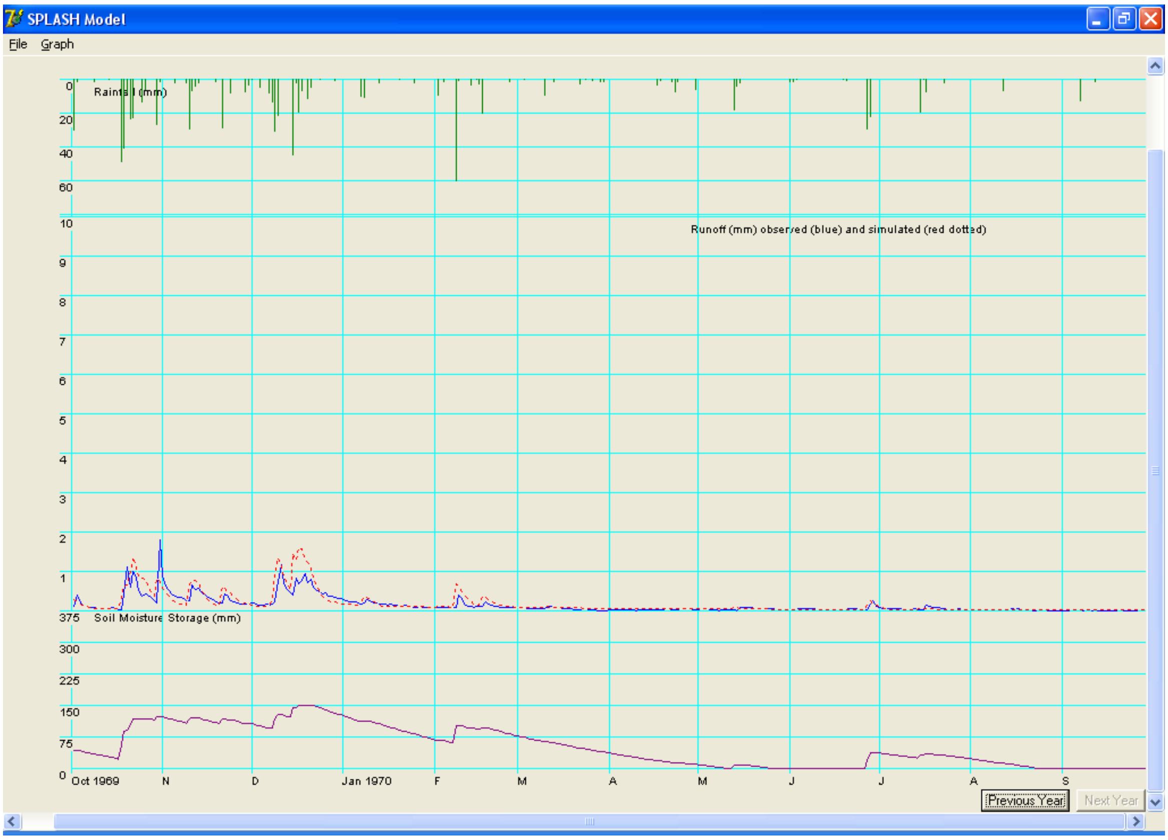


Figure 8: Model performance for the driest year outside the calibration period

8. SPLASH model results for a large catchment

The performance of the model was also tested for a larger catchment with significantly lower rainfall. The selected catchment is not in pristine condition; currently few large catchments are in their natural state. Fortunately long term runoff and rainfall data are available for the Vaal River at Standerton. The mean annual rainfall varies between 750 mm in the east and drops to about 600 mm in the west. The upper reaches of the catchment drain part of the western slopes of the Drakensberg mountain range. The catchment area is 8193 km^2.

Extensive coal deposits are mined within the catchment and provide the raw material for electrical power generation and a fuel from coal plant. The industries required more water which necessitated the building of the Grootdraai Dam as well as inter basin water transfers. The latter developments were commissioned subsequent to 1976 while the runoff record from October 1970 until September 1976 was used for calibrating the SPLASH model and the period 1961/62 to 1969 /70 used for model validation.

Cattle and dry land farming are the main agricultural activities in the catchment. An appreciable number of small farm dams were built for irrigation, stock watering, erosion protection and ground water recharge. No correction for agricultural water use has been made to the flow record due to a paucity of historical water use data. Rather than simulating runoff for normal operational or design purposes, the focus on calibration of the model centred on:
(a) high flows as a means towards enhancing flood management of the Grootdraai Dam which was built in the latter half of the seventies just upstream of the gauging station and
(b) testing of the SPLASH model in a different hydrological regime to that in which it was conceived.

As expected the model did not perform well for low flows due to the unknown water usage in the catchment. Figure 9 illustrates the high flow simulation of runoff with the SPLASH model compared to the observed flow. The daily correlation of simulated versus observed flow for the 1964/65 hydrological year was found to be 0.81.

Figure 9: Vaal River high flow simulation outside of the calibration period. Lower observed peaks since Dec 1964 has unnatural shapes.

9. Conclusion

Although further testing and model refinement should continue, the results obtained with the SPLASH model seem inspiring for South African conditions. Time and data availability from the SADC Region (outside South Africa) prevented wider testing of the model. However, it is expected that the SPLASH model would perform equally well in similar hydrological regimes elsewhere.

Results with the SPLASH model displayed very clear distinction of wet and dry spells. This inspires confidence for model application in both flood forecasting and generation of low flow outcomes based on long or medium term weather forecasts.

Due to the daily update of the lower zone storage and the reliance of simulated flows on the lower zone, the antecedent conditions should enable the SPLAH model to yield better results for infilling of missing daily flow records than statistical techniques such as multiple linear regression or black-box type of approaches. Successful calibration of the model provides for feasible extension of flow records provided longer period catchment rainfall is available.

Ideally the SPLASH model should be combined with routing procedures or models to simulate runoff for very large catchments. This should be done by runoff modelling of tributaries and routing the results downstream to the point of interest.

It should be remembered that the South African Weather Service (SAWS) rainfall records are reported as total daily rainfall at 08:00 each morning. This automatically introduces a lag of 8 hours between streamflow and rainfall because daily runoff is calculated as an average flow for 24 hours ending at 24:00 hours. The model should be enhanced to include a LAG parameter to best match rainfall and runoff.

Further calibration and use of the SPLASH model should yield some insight regarding model parameters in relation to catchment size, catchment- and river slope, mean annual rainfall, vegetal cover, *etc*. Successful regionalization of model parameters would enable use of the model for simulating runoff for ungauged catchments or transferring modelled results up or downstream to points of interest.

10. Acknowledgement

The support of the Water Division, SADC Secretariat and the co-operating partners of the SADC-HYCOS Phase II Project (European Union, the Netherlands Government, WMO and the Department of Water Affairs and Forestry) to use a small part of my time in development of the SPLASH model and participation at this forum is highly appreciated. Data provided by the South African Weather Service and the Hydrological Services of the Department of Water Affairs and Forestry is gratefully acknowledged.

11. References

Chow V.T. Handbook of applied hydrology (edited Chow V.T.) McGraw Hill Book Co., New York, 1964.

Crawford N.H. and Linsley R.K. Digital simulation in hydrology: Stanford Watershed Model IV. Technical report no. 39, Dept. of Civil Engineering, Stanford University, California, USA. July 1966.

Linsley R.K. Rainfall-runoff models, an overview. In Rainfall-runoff relationship, edited by Singh V.P. Water Resources Publications, Colorado, USA.

Lynch S.D., Dent M.C. Estimating areal mean monthly time series of rainfall. V S.A. National Hydrological Simp., Stellenbosch, 1991.

Nielsen S.A., Hansen E. Numerical simulation of the rainfall-runoff process on a daily basis. Nordic Hydrology 4, p 171-190, 1973.

Pitman W.V. A mathematical model for generating monthly river flows from meteorological data in South Africa. Report 2/73, Hydrological Research Unit, University of the Witwatersrand. 1973.

Pitman W.V., Kakebeeke J.P. WRSM90 Water resources simulation model. Version 2.1 User's Guide, 1991.

Refsgaard J.C., Havnoe K. A mathematical modelling system for flood forecasting. Presented at Int. Conf. on Hydraulic Aspects of Floods and Flood Control, London, September 1983.

Seed A.W. The generation of a spatially distributed daily rainfall database for various weather modification scenarios. WRC Report No. 373/1/92, 1992.

Shulze R.E. ACRU: Background concepts and theory. WRC Report No. 154/1/89, 1989.

Song Z., James L.D. An objective test for hydrologic scale. Water Resources Bulletin, AWRA, Vol. 28 No 5, Sept/Oct 1992.