Sequential Conceptual Simplification of the Effective Rainfall Component of a Rainfall-Streamflow Model for a Small Kenyan Catchment

I.G. Littlewood
Centre for Ecology and Hydrology, Wallingford, Oxon, OX10 8BB, United Kingdom (igl@ceh.ac.uk)

Abstract: A six-parameter soil-vegetation-atmosphere-transfer (SVAT) model, calibrated independently by other researchers using neutron probe soil moisture observations, is applied in series with an efficient three-parameter unit hydrograph (UH) in order to simulate continuous daily streamflow for a small catchment in Kenya. The resultant nine-parameter SVAT-UH model is applied to periods of record not used for its calibration to show that it is a reasonable characterisation of the catchment. An experiment was undertaken assuming that only rainfall, streamflow and air temperature (or potential evaporation) data were available for calibrating this structure of rainfall-streamflow model for the Kenyan catchment. Sequentially reducing the structural complexity of the SVAT module, and consequently the number of its parameters from six to three, while leaving the structure of the UH part unchanged, was not accompanied by a decrease in model calibration performance. This indicates the low level of process complexity in SVAT schemes that can be identified when only rainfall, streamflow and evaporation surrogate data are available for model calibration. However, the efficacy of the independently calibrated six-parameter SVAT module when included in a rainfall-streamflow model is demonstrated for the Kenyan catchment, indicating the value of additional soil moisture measurements for rainfall-streamflow modelling.

Keywords: SVAT models; Unit hydrograph; Streamflow modelling.

1. INTRODUCTION

Systematic measurements of soil moisture and evaporation to assist with rainfall-streamflow modelling are rarely undertaken outside small research catchments and study areas. The more common, but by no means guaranteed, situation is that only rainfall, streamflow and air temperature data from national hydrometric networks are readily available. Catchment-scale continuous simulation rainfall-streamflow models that can make efficient and reliable use of these available network data are therefore of great interest, especially for operational and applied research purposes (e.g. regionalization studies where flows at ungauged sites are modelled).

Physically-based, deterministic, models that need large amounts of data at high spatial densities and high temporal frequencies, and which have a large number of parameters (e.g. greater than 12), are usually too costly to calibrate and operate, especially when many catchments are being modelled. Recourse is commonly made, therefore, to so-called conceptual models that (i) need only readily available data (rainfall, streamflow and air temperature or some other surrogate for evaporation), (ii) do not need detailed descriptions of the catchment (soils, vegetation, topography, etc.) and (iii) retain some semblance of reality by representing dominant catchment-scale processes using a simple model structure and a small number of parameters (parametric parsimony).

Conceptual rainfall-streamflow models for catchments largely unaffected by transfers of groundwater with neighbouring areas usually have two main modules. The first module converts rainfall into effective rainfall (i.e. the unobservable proportion of rainfall which eventually produces streamflow). The second module converts the effective rainfall output from the first module into estimated streamflow. Conceptual models themselves can have many parameters but problems with their meaningful identification are well known. For example, there can be large uncertainties in, and cross-correlations between, the model parameters. Conceptual models with many parameters tend to fit better over their calibration periods than models with fewer parameters but they tend to perform worse over...
non-calibration periods [e.g. Perrin et al. 2001], when they are most needed.

The detailed structures of effective rainfall and streamflow production modules vary between models. Although streamflow comprises contributions of flow from many processes and pathways, it is widely recognized that conceptual models calibrated using only rainfall, streamflow and evaporation surrogate data can represent only a small number of dominant flow components. This paper uses a three-parameter unit hydrograph approach in which the estimated streamflow hydrograph can be separated into its dominant quick flow and slow flow responses [Jakeman et al., 1990; Littlewood and Jakeman, 1991, 1994].

In the 1960s, rainfall, streamflow, potential evaporation and soil moisture were monitored in a number of small catchments in the Aberdare mountains of central Kenya for an intensive study of the water-balance effects of different land-uses [Pereira et al., 1962]. The high-quality data from that study can now be re-analysed using models and modelling techniques not available at the time the data were collected. This paper simulates continuous daily streamflow for the Kimakia ‘C’ catchment, 1958 to 1974, using a soil-vegetation-atmosphere-transfer (SVAT) module [Roberts and Harding, 1996] coupled with the parametrically efficient unit hydrograph module referred to above.

2. THE KIMAKIA ‘C’ CATCHMENT

The Kimakia ‘C’ catchment (0.65 km$^2$) is situated approximately 100 km north of Nairobi (0.5 °S of the equator), at an altitude of about 2440 m. Over the study period of January 1958 to June 1974 the dominant vegetation was indigenous bamboo (Arundinaria alpina) with scattered evergreen forest species. The catchment was instrumented for daily streamflow, basin rainfall and other hydrometeorological measurements [Blackie, et al., 1979]. The hydrometeorological measurements were taken manually twice a day close to the gauging station, and used to calculate potential evaporation for short grass [Penman, 1948]. Soil moisture measurements were made by the neutron probe method [Bell and McCulloch, 1966] at six sites under bamboo vegetation. Further details of the measurement techniques employed are given by Blackie et al. [1979].

Average annual rainfall between 1958 and 1973 was about 2200 mm, occurring mainly in April, May, October and November. Catchment average annual runoff over the same period was 1165 mm (53% of rainfall). Groundwater flux beneath the topographic catchment boundary is not an important component of the water balance. Average annual evapotranspiration, calculated as the difference between rainfall and runoff, was therefore 1035 mm. The deep soils have high hydraulic conductivities and play an important role in modulating variations in streamflow. Rainfall can be seasonally intense (daily totals >50mm are fairly common) but infiltration rates are high and low streamflows are sustained during dry seasons (Figure 1). Streams in the region are the basic source of water supply to smallholdings in the area and for the city of Nairobi.

Figure 1. Kimakia ‘C’ 1st January 1958 to 30th June 1974 (a) daily rainfall, (b) daily streamflow

3. THE RAINFALL-STREAMFLOW MODEL

Roberts and Harding [1996] estimated daily soil moisture deficit (SMD$^d$) for bamboo vegetation using equations (1) to (4) below. They optimised the six parameters $\gamma$, $\delta$, $\beta$, $b$, $A$ and $B$ against measured soil moisture to give the values in Table 1. This model can be classified as a soil-vegetation-atmosphere-transfer (SVAT) model.

Table 1. SVAT model #1 parameters

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\delta$</th>
<th>$\beta$</th>
<th>$b$</th>
<th>$A$</th>
<th>$B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.93</td>
<td>-0.10</td>
<td>0.95</td>
<td>2.0</td>
<td>0.5</td>
<td>-0.004</td>
</tr>
</tbody>
</table>

In equations (1) to (4), $INT$ is interception, $P$ is precipitation, $TRANS$ is transpiration, $PE$ is potential evaporation and $DRAIN$ is drainage (from the base of a conceptual soil column).

For this paper, effective rainfall $u_k$ was calculated by equation (5) given below and input to the unit hydrograph (UH) identification procedure devised by Jakeman et al. [1990], in which estimated streamflow $x_k$ is calculated recursively by equation...
(6). The results for Kimakia ‘C’ presented in this paper were obtained in 1996 (unpublished until now) and prompted Evans and Jakeman [1998] to develop a somewhat similar approach to linking a water balance type of loss module with the unit hydrograph module described by equation (6).

\[
SMD_k = SMD_{k-1} - Pf_k + TRANS_k
+ \text{DRAIN}_k + \text{INT}_k
\]

\[
\text{INT}_k = \gamma (1-\exp(-\delta P_k))
\]

\[
TRANS_k = \beta PE_k \left(1 - \frac{1.5P_k}{24}\right),
\]

where \(0 < P_k < \frac{24b}{1.5}\)

\[
\text{TRANS}_k = 0, \text{ otherwise}
\]

\[
\text{DRAIN}_k = A \exp(-B \cdot SMD_k)
\]

\[
u_k = \frac{-a_1 x_k - a_2 x_{k-2} + b_0 u_{k-\lambda} + b_1 u_{k-2-\lambda}}{\lambda}
\]

\[
\nu_k = \frac{-a_1 x_k - a_2 x_{k-2} + b_0 u_{k-\lambda} + b_1 u_{k-2-\lambda}}{\lambda}
\]

\[
x_k = \nu_k x_{k-1} - \alpha x_{k-2} + \beta u_{k-\lambda} + \gamma u_{k-2-\lambda}
\]

Unit hydrograph dynamic response characteristics (DRCs) for the dominant quick and slow response components of streamflow are given by equations (10) to (15) given in Table 2.

Effective rainfall from 4th October 1968 to 18th November 1971, derived by equations (1) to (5) with parameters as in Table 1 (i.e. the catchment was assumed to be 100% bamboo), gave the unit hydrograph DRCs and model-fit statistics for Kimakia ‘C’ in Table 3. In Table 3: D is the percentage of the variance in observed daily streamflow, \(y_k\), accounted for by SVAT-UH model #1 (89.4%); M is a measure of the water balance between effective rainfall and \(\Sigma y_k\) over the calibration period (100% x (\(\Sigma y_k - \Sigma u_k\))/\(\Sigma u_k\)) (-2.9%); and AP is an indicative average precision of the UH a and b parameters in equation (6) (0.03%).

The UH identification procedure involves a trade-off between a high value for D and a low value for AP [Jakeman et al., 1990]. Over its calibration period the nine-parameter SVAT-UH model #1 gave the model-fit shown in Figure 2a. When applied in simulation-mode over the much longer period 24th February 1959 to 30th June 1974, model #1 performed well over sub-periods not used for its calibration, as shown in Figure 2b (D was 85.4% over the whole simulation period).

Because the calibration period of almost three years was selected to start and finish at approximately the same level of low flow (Figure 2a), the volume of modelled effective rainfall should match that of

<table>
<thead>
<tr>
<th>DRC</th>
<th>Quick flow</th>
<th>Slow flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau^q)</td>
<td>(\tau^q = \frac{-\Delta}{\ln(-a_1)})</td>
<td>(\tau^q = \frac{-\Delta}{\ln(-a_1)})</td>
</tr>
<tr>
<td>(\nu^q)</td>
<td>(\nu^q = \frac{-b^q}{1+a^q}\left[1 + \frac{a^q}{\nu}\right]^2)</td>
<td>(\nu^q = \frac{-b^q}{1+a^q}\left[1 + \frac{a^q}{\nu}\right]^2)</td>
</tr>
<tr>
<td>(\rho^q)</td>
<td>(\rho^q = \frac{-b^q}{b^q + b^q})</td>
<td>(\rho^q = \frac{-b^q}{b^q + b^q})</td>
</tr>
</tbody>
</table>

\(\tau^q\) : Characteristic decay response times, where \(\Delta\) = data time step, e.g. 1 day

\(\nu^q\) : Relative volumetric throughflow, where

\[
V = \frac{b^q}{b^q + b^q} + \frac{b^q}{1 + a^q}
\]

\(\rho^q\) : Relative magnitudes of sub-unit hydrograph peaks

Table 2. Unit hydrograph DRCs

<table>
<thead>
<tr>
<th>DRC</th>
<th>Quick flow</th>
<th>Slow flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau^q)</td>
<td>(\tau^q = \frac{-\Delta}{\ln(-a_1)})</td>
<td>(\tau^q = \frac{-\Delta}{\ln(-a_1)})</td>
</tr>
<tr>
<td>(\nu^q)</td>
<td>(\nu^q = \frac{-b^q}{1+a^q}\left[1 + \frac{a^q}{\nu}\right]^2)</td>
<td>(\nu^q = \frac{-b^q}{1+a^q}\left[1 + \frac{a^q}{\nu}\right]^2)</td>
</tr>
<tr>
<td>(\rho^q)</td>
<td>(\rho^q = \frac{-b^q}{b^q + b^q})</td>
<td>(\rho^q = \frac{-b^q}{b^q + b^q})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3. SVAT-UH model #1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau^q)</td>
</tr>
<tr>
<td>(days)</td>
</tr>
<tr>
<td>7.3</td>
</tr>
</tbody>
</table>
observed streamflow. The Roberts and Harding [1996] SVAT model gave a volume of effective rainfall only 2.9% greater than that of observed streamflow (M = -2.9% in Table 3). This degree of water imbalance over the model calibration period is considered to be acceptable and provides a measure of the utility of the SVAT model presented by Roberts and Harding [1996] when applied as part of a rainfall-streamflow model for Kimakia ‘C’.

### 4. SIMPLIFYING THE SVAT MODULE

An experiment was undertaken to investigate the effect of sequentially reducing the structural complexity of, and the number of parameters in, the SVAT part of the rainfall-streamflow model for Kimakia ‘C’, leaving the structure of the UH part unchanged with three parameters. At each stage of simplification all of the model parameters (SVAT and UH) were re-calibrated. The first conceptual simplification (from six to five SVAT parameters, model #2) was to replace the two-parameter representation of interception given by equation (2) with the one-parameter threshold function given by equation (16).

\[
INT_k = P_k, \quad P_k < \gamma^* \\
INT_k = \gamma^*, \quad \text{otherwise}
\]  

(16)

The second simplification (from five to four SVAT parameters, model #3) was to set \( b \) in equation (3) to unity, effectively reducing the transpiration component of the SVAT module to a single-parameter relationship. The third simplification (from four to three SVAT parameters, model #4) was to set \( \gamma^* \) in equation (16) to zero, effectively removing the interception process from the model. In each case ‘best’ SVAT module parameters were selected such that, when they were applied with an optimal UH module, |M| < 1%. The results of this exercise are given in Tables 4 and 5.

**Figure 2.** Model-fits (a) calibration, (b) simulation

For comparison, the loss module described by equations (17) to (19) [Jakeman and Hornberger, 1993] was applied in series with the same UH module structure used for models #1 to #4 (to give model #5). In equations (17) to (19): \( s_k \) is a catchment wetness index (0 < \( s_k < 1 \) (dimensionless); \( \tau_w \) is a catchment drying time constant (e.g. days) given by the value of \( \tau_w(t) \) at a reference temperature, \( R(\circ C) \); \( f \) is a temperature modulation factor (\( \circ C^{-1} \)); and \( C \) is a volume-forcing constant (\( mm^{-1} \)). In this paper \( R = 0 \circ C \).

The parameters of model #5 referred to above, for which \( D = 84.5\% \) and \( AP = 0.07 \), were derived using PC-IHACRES v1.02 [Littlewood et al., 1997] and are given in Table 6. Several practical aspects of modelling with PC-IHACRES v1.02 are discussed by Littlewood [2001]. When applied in simulation mode from 24\textsuperscript{th} February 1959 to 30\textsuperscript{th} June 1974 model #5 gave a value for D of 80.7\%, i.e. substantially less than the 85.4\% obtained for model #1.

As might be expected, inspection of Tables 3, 4 and 5 reveals that model #1, which uses the SVAT parameters given in Table 1 and where only the three UH parameters were optimised against flow measurements, has a lower value of D (89.4\%) than

<table>
<thead>
<tr>
<th>Model</th>
<th>( \gamma^* )</th>
<th>( \beta )</th>
<th>( b )</th>
<th>( A )</th>
<th>( B )</th>
</tr>
</thead>
<tbody>
<tr>
<td>#2</td>
<td>1.4</td>
<td>0.88</td>
<td>2.0</td>
<td>0.6</td>
<td>-0.004</td>
</tr>
<tr>
<td>#3</td>
<td>1.5</td>
<td>0.95</td>
<td>-</td>
<td>0.5</td>
<td>-0.004</td>
</tr>
<tr>
<td>#4</td>
<td>-</td>
<td>1.3</td>
<td>-</td>
<td>0.5</td>
<td>-0.004</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>( \tau^{w0} ) (days)</th>
<th>( \tau^{s0} ) (days)</th>
<th>( \nu^s )</th>
<th>( D ) (%)</th>
<th>( M ) (%)</th>
<th>( AP ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#2</td>
<td>7.07</td>
<td>77.6</td>
<td>0.21</td>
<td>90.3</td>
<td>-0.2</td>
<td>0.03</td>
</tr>
<tr>
<td>#3</td>
<td>7.17</td>
<td>86.8</td>
<td>0.22</td>
<td>90.0</td>
<td>-0.6</td>
<td>0.03</td>
</tr>
<tr>
<td>#4</td>
<td>6.20</td>
<td>141</td>
<td>0.23</td>
<td>91.8</td>
<td>-0.2</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>( u_k = \tau_w \frac{(s_k + s_{k-1})}{2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>#5</td>
<td>( s_k = C \tau_w \left(1 - \frac{1}{\tau_w(t_k)}\right) s_{k-1} )</td>
</tr>
<tr>
<td>#5</td>
<td>( \tau_w(t_k) = \tau_w \exp(0.062f(R - t_k)) )</td>
</tr>
</tbody>
</table>
models #2 to #4 where all the parameters were optimised against flow (D between 90.0% and 91.8%). Note also that the SVAT module parameters of model #1 were not selected such that the water imbalance |M| over the model calibration period was less than 1%, as were models #2 to #4. The results for models #2 to #4 in Tables 4 and 5 allow speculation that model performance actually increases as the number of SVAT module parameters decreases. The slight decrease in D from model #2 to #3 (90.3% to 90.0%) is unlikely to be significant statistically but it could have occurred partly because M for model #3 was -0.6%, whereas it was -0.2% for models #2 and #4.

It can also be noted from Tables 3, 4 and 5 that, as the number of SVAT module parameters decreases, $t^0$ decreases from 7.3 days to 6.2 days, $\tau$ increases from 66 days to 142 days, and $\nu$ increases from 0.189 to 0.230. Dynamic response characteristic $\nu$ indicates the average proportion of streamflow comprising slow flow, and is a Slow Flow Index analogous to the Base Flow Index [e.g. Gustard et al., 1989], which is 0.61 for Kimakia ‘C’ over the model calibration period. The large discrepancy between SFI ($\nu^0$) and BFI for Kimakia ‘C’ is interesting but cannot be resolved in this paper. For many catchments in the UK, SFI tends to be less than BFI but not by as much as for Kimakia ‘C’.

### Table 6. Model #5 parameters

<table>
<thead>
<tr>
<th>$\tau$ (days)</th>
<th>$f$ (C°)</th>
<th>$1/C$ (mm)</th>
<th>$t^0$ (days)</th>
<th>$\tau^0$ (days)</th>
<th>$\nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>3.2</td>
<td>393</td>
<td>4.55</td>
<td>40.9</td>
<td>0.27</td>
</tr>
</tbody>
</table>

The results for models #2 to #4 in Tables 4 and 5 allow speculation that model performance actually increases as the number of SVAT module parameters decreases. The slight decrease in D from model #2 to #3 (90.3% to 90.0%) is unlikely to be significant statistically but it could have occurred partly because M for model #3 was -0.6%, whereas it was -0.2% for models #2 and #4.

5. CONCLUDING REMARKS

Employing the same calibration period, nine-parameter model #1 comprising the Roberts and Harding [1996] SVAT module and the UH module performs better in calibration mode (D = 89.4%, AP = 0.03) than six-parameter model #5 comprising the Jakeman and Hornberger [1993] loss module with the same UH module (D = 84.5%, AP = 0.07). This could be simply because model #1 has three more parameters than model #5. However, over the period 24th February 1959 to 30th June 1974, model #1 also performs better than model #5 in simulation mode (D = 85.4% for model #1 compared with D = 80.7% for model #5), indicating that model #1 really is the better characterization of the catchment. This result points towards the benefit of having soil moisture measurements, in addition to rainfall, streamflow and evaporation surrogate data, with which to calibrate the parameters of the SVAT component of the rainfall-streamflow model. Given the appeal and proven utility of the concepts of effective rainfall and the unit hydrograph (especially the three-parameter unit hydrograph module employed in this paper), it appears that it is the structure and calibration of SVAT (or loss) modules where effort might be best rewarded in an attempt to make further progress in estimating continuous streamflow using parsimonious conceptual rainfall-streamflow models.

Simple loss modules other than the one given by equations (17) to (19) have been applied with the three-parameter unit hydrograph employed in this paper [e.g. Evans and Jakeman, 1998]. These other loss modules have also indicated that when the available data are restricted to time series of rainfall, streamflow and an evaporation surrogate it is difficult to identify more than three SVAT module parameters. The analysis presented in this paper, where the structure of a six-parameter loss module was sequentially simplified to become a five-, four- and three-parameter loss module without discernible worsening of model-fit performance (indeed, there was a slight indication that model performance actually improved), further supports the notion that three is, in practice, the maximum number of SVAT parameters that can be identified reliably under such restrictions of data availability. However, this paper has also shown that additional soil moisture data may allow reliable identification of more than three SVAT module parameters when they are calibrated within a rainfall-streamflow model.

The parameter identification technique adopted for the analyses reported in this paper does not distinguish explicitly between one part of the record or another regarding its information content for identifying a particular model parameter. For example, it is likely to be difficult to calibrate a loss (or SVAT) module parameter that controls the rate of evaporation when the data being used for that calibration describe a flood event, during which evaporation is of minor (if any) importance in the dynamic water balance. There is growing support for an alternative approach to model parameter identification whereby different parts of the record are used to identify different model parameters [e.g. Wagener and Wheater, 2002]. In this context it is interesting to note that when Roberts and Harding [1996] calibrated the SVAT module given by (1) to (4) they employed different parts of the available record for calibrating different parameters, e.g. the drainage parameters $A$ and $B$ in equation (4) were estimated using only periods when soil moisture was high. This may
contribute to the efficacy of their SVAT module when coupled with the unit hydrograph module, as demonstrated in this paper.

A judicious blend of the use of novel parameter identification techniques and additional soil moisture data (where available) for rainfall-streamflow modelling can be expected, therefore, to lead to either (a) more than three SVAT (or loss) module parameters being identifiable with moderate precision or (b) about three SVAT module parameters being identified with better precision. The choice will depend, as always, on the intended application of the model.

6. REFERENCES


