Hydrol. Earth Syst. Sci. Discuss., 6, 5753–5782, 2009
www.hydrol-earth-syst-sci-discuss.net/6/5753/2009/
© Author(s) 2009. This work is distributed under
the Creative Commons Attribution 3.0 License.

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Selection of an appropriately simple storm runoff model

A. I. J. M. van Dijk

CSIRO Land and Water, Canberra, ACT, Australia

Received: 12 August 2009 – Accepted: 12 August 2009 – Published: 9 September 2009

Correspondence to: A. I. J. M. van Dijk (albert.vandijk@csiro.au)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Alternative conceptual storm runoff models, including several published ones, were evaluated against storm flow time series for 260 catchments in Australia (23–1902 km^2). The original daily streamflow data was separated into baseflow and storm flow components and from these, event rainfall and storm flow totals were estimated. For each tested model structure, the number of free parameters was reduced in stages. The appropriate balance between simplicity and explanatory power was decided based on Aikake's Final Prediction Error Criterion and evidence of parameter equivalence. The majority of catchments showed storm recession half-times in the order of a day, with more rapid drainage in dry catchments. Overland and channel travel time did not appear to be an important driver of storm flow recession. A storm runoff model with two free parameters (one related to storm event size, the other to antecedent baseflow) and a fixed initial loss of 12 mm provided the optimal model structure. The optimal model had some features similar to the Soil Conservation Service Curve Number technique, but performed an average 12 to 19% better. The non-linear relationship between event rainfall and event runoff may be associated with saturated area expansion during storms and/or the relationship between storm event size and peak rainfall intensity. Antecedent baseflow was a strong predictor of runoff response. A simple conceptual relationship between groundwater storage and saturated catchment area proved adequate and produced realistic estimates of saturated area of $<0.1\%$ for the driest and $>5\%$ for the wettest catchments.

1 Introduction

1.1 Rationale

Predicting storm rainfall-runoff response is an important component of catchment hydrology. Experimental studies have shown that storm runoff occurs through a variable combination of Hortonian overland flow (rainfall intensity exceeds soil infiltration capacity), saturation (or Dunne) overland flow (rainfall onto saturated and perhaps frozen soil), subsurface storm flow (a lateral subsurface flow component often associated with an impeding layer) and direct rainfall onto waterways, water bodies and riparian areas saturated for longer periods of time. A combination of these mechanisms is likely in many cases, with infiltration excess runoff dominating in dry climates with intensive rainfall (e.g. semi-arid tropics) and the other mechanisms in more humid environments. It follows that explaining observed rainfall-runoff response (and by extension estimating it where it is not observed) can be challenging, particularly as knowledge of the distribution of rainfall in time and space is limited by the number of gauges and the (still often daily) time step of rainfall data.

1.2 Existing model formulations

A considerable number of conceptual models have been developed over the years to estimate storm flow from daily rainfall totals (reviewed in e.g. Beven, 2004; Blöschl, 2005; Maidment, 1992). Most models recognise the apparent threshold behaviour or non-linearity in the relationship between event runoff and event (or more commonly, daily) rainfall. Alternative processes can be assumed responsible for the threshold or non-linearity; for example, it may be considered that larger storms: (1) tend to have higher peak rainfall intensities; (2) are more likely to temporarily saturate soil horizons; (3) are less affected by initial losses such as canopy wetting; (4) fill soil moisture and/or groundwater stores, forcing out infiltrated rainfall; or (5) fill surface retention stores (e.g. hollows, small dams). Threshold and non-linear processes may well lead to an identical type of response when it is considered that thresholds associated with terrain or vegetation properties have spatial variability, and that rainfall has both temporal as well as spatial variability. Spatial variability in water inputs are further enhanced by the influence of vegetation on below-canopy precipitation distribution and by run-off–run-on processes.

Different models conceptualise rainfall-runoff response differently and make different (often implicit) assumptions about dominating runoff processes. Many of the simpler models that can be used with daily time step rainfall data use simple threshold functions. A generic set of equations that captures most thresholds and variables usually considered may be:

$$R = f_{\mathrm{sat}} P_n + (1 - f_{\mathrm{sat}}) P_n - I + R_{\mathrm{return}} \quad (1)$$

$$P_n = \max(0, P - L_{\mathrm{i}}) = (1 - f_n) P \quad (2)$$

$$I = f_I P_n \quad (3)$$

$$R_{\mathrm{return}} = (1 - f_s) I \quad (4)$$

where P_n net rainfall, R is runoff, I net infiltration into the soil, R_{return} return flow from the soil (all in mm per event or mm d^{-1}), and f_{sat} the fraction saturated area. P_n is often expressed as total rainfall minus an initial loss L_{i} (mm) which may be conceptualised as a constant or a proportion f_n (or a combination of both) and assumed to represent an infiltration loss or an evaporative loss. The fraction f_I represents the fraction of net rainfall on unsaturated soil that infiltrates, and f_s the fraction of I that can be retained in the soil. This generalised model is usually simplified one way or another, for example, by assuming that f_{sat}, f_n, f_I or f_s are either negligible or equal to unity (e.g. Bergström, 1992; Chiew et al., 2002). It may also be made more complicated, by introducing further functional relationships, for example expressing f_{sat} as a function of groundwater level or storage, or expressing f_I as a function of storm size or rainfall intensity and/or soil water content, or expressing f_{soil} as a function of actual and maximum soil water storage. These functions will introduce further free parameters, at least if the assumed relationship is not simply directly proportional. Additionally, one or several of the variables in these equations may be represented by spatial (non-linear) distribution functions, for example where required to represent sub-grid variability in coarse resolution land surface models (e.g. Bonan, 1996; Liang et al., 1994; Liang and Xie, 2001; Oleson et al., 2004).

Even in its relatively simple basic form, Eq. (1) and its constituent Eqs. (2)–(4) have considerable scope for equivalent parameterisations, and therefore for practical use of this set of equations it either needs to be possible to estimate parameter values with confidence, or simplifications need to be introduced. The well-established Soil Conservation Service Curve Number method (SCS-CN; USDA-SCS, 1985) can be taken as an example of the latter. It can be recast in a form somewhat similar to Eq. (1) as:

$$R = P_n \left(\frac{P_n}{P_n + S} \right) \tag{5}$$

Where S is notionally the maximum retention after runoff begins (mm). The SCS-CN model was empirically derived however, and when comparing to Eqs. (1)–(4) the second term can be interpreted in more than one way:

- Equivalent to $(1-f_n)$, representing the functional relationship between storm size and fraction rainfall in excess of infiltration capacity, in which case P_n could be interpreted as a proxy for rainfall intensity and S as a proxy for maximum infiltration rate (while f_{sat}=0 and f_s=0).
- Equivalent to $(1-f_s)$, representing the functional relationship between storm size and return flow fraction, in which case S represents maximum soil storage capacity (while f_{sat}=0 and f_l=1).
- Equivalent to f_{sat}, representing a functional relationship between storm size and saturated catchment area that could arise if cumulative infiltration or run-off-run-on processes lead to increase of the saturated area over the course of a storm, where P_n might be a proxy for cumulative infiltration or actual runoff, and S a proxy of the efficiency of soil and catchment drainage (in which case f_n=1 and f_s=0).

Elaborations of the SCS-CN method that implicitly or explicitly assume one of these three underlying explanations (or a combination) do exist, by relating S to land cover or soil conditions and/or introducing dynamic behaviour by modifying S as a function of antecedent rainfall or groundwater storage (for examples of both see Maidment, 1992; Mishra and Singh, 2003). It would be expected that the effectiveness of these elaborations will depend on the dominant runoff processes.

1.3 Objective

To our knowledge, there has not been a comprehensive analysis to assess which of several alternative model formulations best explains the variation in rainfall-runoff response observed in catchment streamflow data. For example, common approaches to runoff estimation in Australia assume a relationship between runoff response (that is, storm flow as a fraction of rainfall) and event size (e.g. using the SCS-CN method) or relative soil water content, yet previous analyses of streamflow observations in Australian catchments suggested that runoff response increases during the course of the wet season in phase with baseflow, rather than responding to (shallow) soil moisture or event size only (e.g. Liu et al., 2007; Peña Arancibia et al., 2007; E. Kwantes, unpublished). This is illustrated for two catchments in Fig. 1. In this study, we have the following aims:

- Test several alternative versions and simplifications of the generalised storm flow models listed for their performance in reproducing estimated event storm flow from 260 catchments across Australia.
- Evaluate to what extent incorporation of (estimated) catchment groundwater storage into the model structure can improve event storm flow prediction.
- Assess the appropriate balance between the number of fitting parameters and model performance, by considering an information criterion as well as correlation between different parameters that may be indicative of equivalence.
- Based on this assessment, select an appropriately simple storm runoff model.

2 Methods

2.1 Data

Daily streamflow data (in ML d^{-1}) were collated for 362 catchments across Australia as part of previous studies (Guerschman et al., 2008; Peel et al., 2000). The catchments where selected to have streamflow data of satisfactory quality and any influence of river regulation, water extraction, urban development, or other processes upstream streamflow was considered unimportant. The contributing catchments of all gauges were delineated through digital elevation model analysis with visual quality control. Catchment areas vary between 23–1979 km^2, with a median value of 315 km^2. The range of average annual rainfall for the catchment sample was 3177 to 1887 mm y^{-1}, whereas Priestley-Taylor potential evapotranspiration (E_0) varied from 765 to 2417 mm y^{-1} and total streamflow from 2 to 979 mm y^{-1}. Catchment humidity (H, the ratio of average rainfall over average E_0) varied from 0.13 to 3.48.

From the data set, those records were selected that had good quality observations for at least five years during the period 1990–2006 and no less than 50 runoff events (defined as an increase in streamflow from one day to the next). The streamflow data for these 260 catchments were converted to areal average streamflow (Q, mm d^{-1}); their location is shown in Fig. 2.

2.2 Streamflow analysis

The streamflow data were separated into time series of daily estimated baseflow (BF or Q_{BF}) and quick flow (QF or Q_{QF}). It was assumed that BF represents groundwater discharge only, and QF the sum of all storm runoff processes, but it is noted that the hydrograph per se does not provide evidence for this interpretation. The BF and QF components were estimated from the streamflow record by combining forward and backward recursive linear reservoir baseflow filter. The approach follows that of Wittenberg (1999) and is summarised in Appendix A.

To estimate the storm flow recession constant (k_{QF}) from the estimated storm flow time series for each station, all days with a storm flow peak (that is, $Q_{QF}(t)>Q_{QF}(t-1)$ and $Q_{QF}(t)>Q_{QF}(t+1)$) were identified. A weighted-average storm flow recession constant was calculated from the storm flow estimates for these (denoted t_*) and the subsequent (t_*+1) days as:

$$k_{QF} = 1 - \frac{\sum Q_{QF}(t_* + 1)}{\sum Q_{QF}(t_*)}. \tag{6}$$

For each runoff event i (identified by an increase in q_{QF} on day t following a decrease or no change on day $t-1$) total event runoff R and rainfall P were estimated in one of two ways. If $Q_{QF}(t)>Q_{QF}(t+1)$:

$$R(i) = \frac{Q_{QF}(t)}{k_{QF}} \quad \text{and} \quad P(i) = P(t-1) + P(t) \tag{7}$$

Otherwise, if $Q_{QF}(t)\leq Q_{QF}(t+1)$:

$$R(i) = \frac{Q_{QF}(t) + Q_{QF}(t+1)}{k_{QF}} \quad \text{and} \quad P(i) = P(t-1) + P(t) + P(t+1) \tag{8}$$

The second case occurred where the majority of storm runoff passed the gauge a day after the majority of precipitation had fallen. Progressive further increases in streamflow beyond day $t+1$ were observed for a very small number of events; these were attributed to rainfall or streamflow measurement issues and the associated events omitted from the analysis.

Groundwater storage S_g was estimated from the daily baseflow (Q_{BF}) estimates as:

$$S_g(t) = \frac{Q_{BF}(t)}{k_{BF}} \tag{9}$$

where k_{BF} is the baseflow recession constant. Values of S_g on the first day of each runoff event were used as an estimate of antecedent groundwater storage $S_g(i)$ before runoff event i. Due to the method of baseflow separation these values were estimated from the preceding baseflow recession with a forward filter (see Appendix A) and therefore not affected by storm flow event itself.

2.3 Evaluation of alternative model structures

A top-down approach to model selection was followed, that is, the most complex model was gradually simplified and the change in prediction error assessed. The basic most model structure tested had the form:

$$P_n = \max(0, P - C_1) \tag{10}$$

$$R = \left[\frac{(C_2 P_n)^{C3}}{P_n + C_4}\right] P_n \tag{11}$$

where C_1, C_2, C_3 and C_4 are all optionally free parameters. Each of these parameters can be effectively omitted by giving it a value of either zero or unity. For example, C_4=0 simplifies the equation to a power function of P_n, while in addition C_3=1 produces a constant runoff fraction, and C_1=0 removes the rainfall threshold needed to produce runoff. With C_3=1 and C_2=1 the equation mirrors the SCS-CN model. To test whether model performance was further improved if the model took into account the effect of antecedent groundwater storage S_g on saturated area or the available catchment storage (in terms of the SCS-CN method), the following modifications were also tested:

$$R = \left(\frac{(C_2 P_n)^{C3}}{P_n + C_4}(1 - f_{\text{sat}}) + f_{\text{sat}}\right) P_n \quad \text{with} \quad f_{\text{sat}} = \max\left(1, C_5 S_g^{C6}\right) \tag{12}$$

and

$$R = \left[\frac{(C_2 P_n)^{C3}}{P_n + S_{\max}}\right] P_n \quad \text{with} \quad S_{\max} = C_4 \min\left(0, 1 - C_5 S_g^{C6}\right), \tag{13}$$

where C_5 and C_6 are again parameters that could be fitted or prescribed values of zero or unity. It follows that the most complex models had six free parameters.

To assess the trade-off between the number of free model parameters and the improvement in model performance, Akaike's Final Prediction Error Criterion (FPEC; Akaike, 1970) was used. FPEC has been widely used in model selection and represents the expected prediction error that would result were the model tested on a different data set. FPEC is calculated as the product of an empirically estimated prediction error (ε) and a penalization factor that considers the degrees of freedom (d, the number of free parameters) in comparison to the number of observations (n):

$$\text{FPEC} = \frac{1 + d/n}{1 - d/n}\varepsilon \tag{14}$$

Volumetric estimation error was chosen as an estimate of ε (Nash–Sutcliffe efficiency or similar indicators based on squared difference were avoided for reasons outlined by Criss and Winston, 2008; Legates and McCabe, 1999):

$$\varepsilon = \frac{\sum |R_{\text{est}} - R|}{\sum R}. \tag{15}$$

For each of the stations ε was calculated by using it as the objective function for parameter estimation. Latin Hypercube Sampling (performed in MATLAB®) was used to find the optimal parameter set for each model variant. The value of n was calculated as the number of events used to fit the model (599 on average; see results). In principle, the model with the lowest FPEC should be adopted. However, Schoups et al. (2008) pointed out that Eq. (14) assumes that $n \gg d$ and may lead to underestimates of prediction error and favour overly complex models. Because most of the model error is likely to be associated with a small number of large storm flow events, the calculated FPEC value should probably be considered an optimistic estimate. This caveat was considered when interpreting FPEC values, and other factors assessed were: (1) the number of stations for which alternate model structures appeared to function best; and

CC BY

(2) the correlation between fitted parameters. A parametric or non-parametric (ranked) coefficient of correlation (r and r_*, respectively) between fitted parameters of 0.40 or more was considered an indicator of possible equivalence between model parameters.

3 Results

3.1 Storm flow recession coefficient

The distribution of storm flow recession coefficients (k_{QF}) appeared more or less normal with an average of 0.52 (st. dev. ±0.11) and half of all values between 0.45–0.60 (Table 1). It follows that the storm flow decay period (or "half time") was generally about a day. The highest k_{QF} values (i.e. fastest recessions) occurred in dry catchments: catchment humidity index H showed some correlation with k_{QF} ($r_*=-0.49$) and a linear regression explained 21% of the variance.

3.2 Optimal model structure

The average station record included 1902 storms (range 679–3454) of which 32% (range 2–58%) produced storm flow, resulting in an average 599 events per station (range 30–967). Equation (14) indicates that the average improvement in the error estimate therefore would need to be 0.5% or more for an additional parameter to be accepted. For 18 out of a subset of 20 stations, fitting the five and six parameter models for a subset of stations produced worse FPEC scores than fitting the four parameter model. Computational time to assess all six variants (5 five-parameter models and one six-parameter model) was very considerable, because the larger number of parameters that required simultaneous optimisation. To avoid this, only all model variants with four or less parameters were optimised for all 260 stations.

The four-parameter model structure with the best FPEC score was Eq. (12) with $C_1=1$ and $C_3=1$. Alternative models with the exponent C_3 as a free parameter produced an apparently normal distribution of values with mean 1.10 (st. dev. ±0.38), but did not lead to a better FPEC score. The distribution of FPEC values for all stations as the number of free parameters is reduced is shown in Fig. 3. These reflect the best results obtainable, and were achieved by (from left to right):

- Reducing from four to three parameters by fixing C_6. The fitted values appeared normally distributed, with an average of 0.94±0.56 (Table 1). Prescribing C_6 to unity increased FPEC for 51% of stations and increased average FPEC by 1.6% compared to the four-parameter model.
- Reducing from three to two parameters by fixing C_2. The distribution of fitted C_2 values was positively skewed and had a median value of 11.6 mm (Table 1). Prescribing this value to C_2 increased FPEC for 77% of stations and increased average FPEC by 3.2% compared to the four-parameter model.
- Reducing from two to one parameter by fixing C_4. The distribution of fitted C_4 values was strongly positively skewed and had a median value of 4667 mm (Table 1). Prescribing C_4 to this value increased FPEC for 83% of stations and increased average FPEC by 5.5% compared to the four-parameter model.
- Removing all free parameters by fixing C_5. The distribution of fitted C_5 values was positively skewed and had a median value of 0.0048 mm^{-1} (Table 1). Prescribing C_5 to this value increased FPEC for 95% of stations and increased average FPEC by 14% compared to the four-parameter model.

3.3 Parameter values and correlation

The dimensional (absolute) volumetric error was calculated by multiplying FPEC with average QF. It appears that errors were directly proportional to average QF (r^2=0.98), and that FPEC itself was correlated to the number of rain days: the 20 dry catchments with the most erratic rainfall had relative errors that were on average ca. 40% higher than the 20 humid catchments with the most regular rainfall.

HESSD
6, 5753–5782, 2009

CC BY

Optimised values for parameter C_2 in the three-parameter model (interpreted as initial loss L_i in mm in the remainder) showed a positively skewed distribution, with a median value of 11.6 mm and 80% of values between 5–48 mm (Table 1). Values for parameter C_4 in the three-parameter model (interpreted as maximum storage capacity S_{max} in mm) had an approximately log-normal distribution with a median value of 5667 mm, and 80% of values between 5–48 mm). Regression analysis showed that S_{max} values were correlated with catchment storm runoff fraction (period average storm flow over average rainfall) (r_*=−0.50). Values for parameter C_5 in the three-parameter model (denoted the saturated area coefficient C_{Sg} in mm^{-1}) had a strongly skewed distribution. Fitted values showed correlation with catchment climate indices such as period average rainfall ($r^* = -0.57$) and catchment humidity (H, r_*=−0.57). The correlation implies that for a given mm change in catchment groundwater storage, saturated area changes faster in a drier catchment than in a wetter catchment. Finally, some correlation appeared to occur between optimised values of L_i and S_{max} (r_*=−0.42), but not between C_{Sg} and L_i (r_*=0.11) or S_{max} (r_*=0.21).

4 Discussion

4.1 Storm flow recession coefficient

Most catchments showed storm flow recession "half times" of about one day (k_{QF}=0.52 d^{-1}; CV±21%) with the most rapid drainage in dry catchments. This is consistent with the expectation that storm flow under dry conditions would be predominantly through infiltration excess overland flow during a small number of high intensity rainfall events. The 260 catchments varied considerably in size (23 to 1902 km^2) and, because of differences in runoff travel time, therefore a relationship with catchment size might have been expected. Regression analysis did not indicate any such relationship (r=0.21). Travel times estimated with published generalised methods such as those in Maidment (1992) suggested travel times between <0.05 day to ca. 0.5 day for the

HESSD
6, 5753–5782, 2009

Selection of an appropriately simple storm runoff model

A. I. J. M. van Dijk

catchment size range, and ca. 0.1 days for the median 333 km^2 catchment. Compared to the derived recession half times of around one day these numbers are small, and therefore it appears overland and channel storm flow routing is not the main driver of the observed storm flow recessions. It is concluded that storm flow recession is dominated by the release of water that is temporary retained in ephemeral water bodies, draining soil or perched groundwater.

4.2 Optimal storm runoff model structure

Based on the calculated FPEC values, the four-parameter model could be accepted as (theoretically) having the smallest prediction error. However, the marginal deterioration in performance between the four- and three-parameter model (a 1.6% increase in average FPEC and much less for the majority of catchments) was considered to be preferable over the increased likelihood of parameter equivalence problems, particular since the criterion likely underestimated prediction error of the more complex models (for reasons discussed earlier).

Both the four and the three parameter version of Eq. (12) can be changed to a perhaps more insightful notation of parameters:

$$R = \left(\frac{P_n}{P_n + S_{\max}}(1 - f_{\text{sat}}) + f_{\text{sat}}\right) P_n \tag{16a}$$

with

$$P_n = \max(0, P - L_{\text{i}}) \tag{16b}$$

and

$$F_{\text{sat}} = \max(1, C_{\text{Sg}} S_g) \tag{16c}$$

Note that the definition of the maximum water retention capacity ($S_{\max}$) here is similar but slightly different from the parameter S in the SCS-CN model (Eq. 5).

4.3 Comparison to SCS-CN technique

Because of the similarity of the optimal model structure to the SCS-CN model, it is of interest to make a direct comparison of performance. Following the same approach as for the other model structures, the SCS-CN model was fitted in three alternative recommended ways:

1. The conventional two parameter version, optimising I_a (in SCS-CN notation, equivalent to L_i) and S (equivalent to S_{max}) produced a FPEC of 0.81. The values of I_a and S were negatively correlated rather than the positive correlation expected (cf. Maidment, 1992; Mishra and Singh, 2003; USDA-SCS, 1985). Moreover, for 207 out of 260 stations, optimised S values were outside the range of 0–593 mm that corresponds to the recommended range of curve numbers (30–100).

2. The recommended one-parameter version, which assumes $I_a=0.2 \cdot S$ produced a FPEC of 0.88. Estimated S values were within the recommended range for 246 out of 260 stations. Converting the S values fitted to curve numbers suggested an average CN of 46±12.

3. A modified two-parameter version, using the same assumption but further including the extension with f_{sat} and associated parameter C_{Sg} as per Eq. (16) produced a FPEC of 0.87.

It is concluded that the optimal three-parameter model selected outperforms the SCS-CN method by 12 to 19% (FPEC 0.81–0.88 vs. 0.71). Therefore, at least for Australian conditions, Eq. (16) appears to provide an improvement in storm runoff estimation when compared to the SCS-CN technique or indeed any other runoff model that can be expressed in terms equivalent to Eq. (12) or (13).

HESSD

6, 5753–5782, 2009

Selection of an appropriately simple storm runoff model

A. I. J. M. van Dijk

4.4 Initial loss

The correlation between L_i and S_{max} suggests potential for parameter equivalence and using the median value of 12 m caused little deterioration in model performance. Trialling alternative values for L_i suggested that values of 11–13 mm produced the best FPEC, but performance deteriorated <2% for any value between zero and 20 mm. This is reasonably consistent with the SCS-CN approach; as stated a common assumption in applying the method is that initial abstraction equals 0.2 of maximum retention (USDA-SCS, 1985). Using curve number estimates of 60 to 90 (covering most of the range recommended for forest and grazing land) produces initial abstraction estimates of 6–34 mm. Initial losses are a conceptual water balance component covering a variety of processes, including rainfall retained by vegetation canopy and other surfaces and subsequently evaporated (typically in the order of 1–3 mm; e.g. van Dijk and Bruijnzeel, 2001), losses to wet up a dry soil surface and runoff retained in surface depressions that need to be filled before catchment runoff occurs (technically not an initial loss but in practice likely to be lumped into it due to the model structure (see also van Dijk et al., 2005b). A combined total of ca. 12 mm for all these losses would seem realistic.

4.5 Maximum storage capacity

If the model was simplified by assuming L_i=12 mm, the resulting S_{max} values changed by more than 20% for 60% of stations. It is concluded that this parameter is poorly constrained. Values of S_{max} found through optimisation were generally very high: 97% of stations had fitted values of more than 600 mm, and the maximum bound set at 10^6 mm was still suboptimal for 17% of stations. Such high values make interpretation as a "maximum potential retention" unrealistic, and calls for another interpretation of this functional relationship. When the ratio S_{max} over P_n attains high values, Eq. (16a) approaches the linear relationship:

$$R \approx \left((1 - f_{\text{sat}}) \frac{P_n}{S_{\max}} + f_{\text{sat}} \right) P_n \, . \tag{17}$$

For example, for $S_{\max}$=500 and P_n=50, the difference in calculated R is 10%. Fitting Eq. (17) to the data led to an overall deterioration in FPEC of 1.5%. Although Eq. (16) was preferred for its more realistic limits, it may be more conceptually appropriate to rewrite it in the equivalent form:

$$R = \left((1 - f_{\text{sat}}) \frac{k_P P_n}{k_P P_n + 1} + f_{\text{sat}} \right) P_n \, , \tag{18}$$

where k_P is a constant of proportionality that describes the initial increase in runoff fraction with event precipitation.

More than once mechanism can be invoked to explain why runoff fraction should increase with rainfall event, as is implied by Eq. (18). A rapid increase of temporarily saturated surface area as more rainfall accumulates (e.g. because of a perched water table) provides one possible explanation and has been observed in field studies (Latron and Gallart, 2008; Tanaka, 1992). An alternative explanation is that P_n may function as a surrogate for peak storm rainfall intensity; the key storm characteristic if runoff is dominated by Horton overland flow. Field studies in Indonesia by the author demonstrated: (1) that an effective depth-averaged rainfall intensity (DARI) can be calculated for every storm from short intervals measurements; (2) that for a given site this index has strong predictive power to estimate storm runoff coefficient; and (3) that there appeared to be an approximately linear relationship between storm rainfall depth and DARI. These findings could be combined to produce a theory linking event rainfall and runoff coefficient with a functional form that in fact closely resembles that of Eq. (18) and explained observed runoff from study plots of a range of sizes (<1 to 40 000 m^2; van Dijk et al., 2005b; van Dijk and Bruijnzeel, 2004; van Dijk et al., 2005a). The spatially variable infiltration model underlying the theory was originally developed and validated for sites in Australia and several southeast Asian countries (Yu et al., 1997) while the intra-storm rainfall intensity distribution has been shown equally valid for Australia (Surawski and Yu, 2005). In summary, the relationship between event size and runoff fraction can be explained by expansion of the saturated area during the storm or the statistical relationship between event size and peak intensity, or a combination of both.

4.6 Saturated area coefficient

Antecedent baseflow proved a good predictor of storm runoff response. This would not surprise if saturation overland flow associated with groundwater (or other slowly draining stores) is an important runoff generating mechanism. Previous research worldwide has shown that this mechanism can indeed play an important role (see recent review in Latron and Gallart, 2008).

To assess whether the saturated areas associated with the parameter estimates were realistic, the apparent median fraction of saturated area (f_{sat}) was estimated for each catchment by combining the optimised C_{Sg} value with the groundwater storage (S_g) estimated from the median baseflow rate in the time series. The resulting f_{sat} was less than 1% of the area for 72% of the catchments. Values were positively correlated to catchment humidity ($r=0.66$; Fig. 5). Values greater than 5% were calculated for nine out of 260 catchments. Of these, six catchments experienced more than 1300 mm rainfall per year, and therefore it would not seem unrealistic >5% of the area would be saturated for half of the time. The high values for the three other catchments were associated with very low FPEC scores and considered anomalous.

The value of baseflow in explaining runoff response does not imply that water storage in the unsaturated zone, and in the soil in particular, has no effect on runoff response. Previous analysis using satellite-observed wetness of the top few cm of soil showed little value in predicting runoff response in Australian catchments: effectively, the dynamics in these shallow observations were much more rapid than those observed in surface runoff response, which increased more gradually during the course of the wet season in phase with baseflow (Liu et al., 2007; E. Kwantes, unpublished). It may be

CC BY

that deeper soil moisture still plays a role in determining runoff response, for example by influencing rapid sub-surface pathways that allow infiltrated water to generate return flow. We did not have direct observations or reliable estimates of root zone soil moisture content and therefore were not able to investigate this. In future, field observations or soil water content estimates produced by a hydrological model may help to asses this.

4.7 Sources of uncertainty in runoff estimation

Even the best model explained relatively little of the variance in event storm flow totals: the average prediction error of 0.70 implies that the sum of absolute differences between observations and estimates represents 70% of the total storm flow volume (Fig. 4). The lack of consideration for peak rainfall intensities, as stated before a key driver of infiltration excess runoff, is a likely factor. Peak rainfall intensities are correlated to event rainfall depth, but relationships are not very strong (e.g. r^2=0.62 at a site in West-Java; van Dijk and Bruijnzeel, 2004) and differences of an order of magnitude or more may still occur between storms of equal total depth. In addition, the daily rainfall data used in the current analysis are based on interpolation of rainfall gauge data and the density of the network in Australia is generally rather low, if variable. The catchments with the highest FPEC were found to be dry catchments. Many of these are found in regions with sparse gauging, where convective storms can introduce large positive (negative) biases in average catchment rainfall estimates in interpolation if they are (not) gauged. Current developments in rainfall estimation from ground-based radar and remote sensing should help address both constraints and allow improvements in runoff estimation in future.

5 Conclusions

Streamflow data for 260 Australian catchments were used to evaluate the performance of alternative conceptual storm runoff models and derive a model of appropriate simplicity. The following conclusions are drawn:

1. The majority of catchments showed storm recession half-times in the order of a day, with the most rapid drainage occurring in dry catchments with large rainfall events. Overland and channel travel time did not appear to play an important role in storm flow recession for catchments of the size range investigated (<2000 km^2).

2. A model with two free parameters (one related to storm size, the other to antecedent baseflow) and a fixed initial loss of 12 mm was considered to provide the optimal model structure. Optimising initial loss improved model performance by only 1.6% and appeared to lead to some parameter equivalence issues.

3. The derived model structure was somewhat similar in form to the SCS-CN technique but performed an average 12 to 19% better.

4. Even the best model showed moderate performance in predicting event runoff and the constant of proportionality between event size and runoff fraction appeared highly variable. The non-linear relationship between event rainfall and event runoff was attributed to a combination of saturated area expansion during storms and the statistical (but not very strong) relationship between storm event size and peak rainfall intensity.

5. Groundwater storage (or more precisely, antecedent baseflow) proved a strong predictor of runoff response. A conceptual relationship between groundwater storage and saturated catchment area proved adequate and produced realistic estimates of saturated area with median values of <0.1% for the driest and >5% for the wettest catchments.

HESSD
6, 5753–5782, 2009

Appendix A

Baseflow separation method

Starting at the second last value of the stream flow time series ($i=N-1$) and moving backwards through the record, baseflow for time step i was estimated by considering the forward estimate $Q_{\mathrm{BF},f}$ and backward estimate $Q_{\mathrm{BF},b}$ calculated as:

$$Q_{\mathrm{BF},b}(i) = \frac{1}{1-k_{\mathrm{BF}}} Q_{\mathrm{BF}}(i+1)\,. \tag{A1}$$

The forward estimate of baseflow $Q_{\mathrm{BF},f}(i)$ is given by:

$$Q_{\mathrm{BF},f}(i) = (1-k_{\mathrm{BF}}) Q_{\mathrm{BF}}(i-1)\,, \tag{A2}$$

where $Q(i-1)$ equalled zero, $Q_{\mathrm{BF},f}(i)$ was also given a value of zero. The backward or forward baseflow estimate was assigned using the following decision tree:

1. If $Q(i)<Q(i-1)$ (i.e. falling limb):
   - (a) If $Q_{\mathrm{BF},b}(i)<Q(i)$ then the backwards estimate was adopted: $Q_{\mathrm{BF}}(i)=Q_{\mathrm{BF},b}(i)$
   - (b) If $Q_{\mathrm{BF},b}(i)\geq Q(i)$ then:
     - i. If $Q_{\mathrm{BF},f}(i)<Q(i)$ the forward estimate was adopted: $Q_{\mathrm{BF}}(i)=Q_{\mathrm{BF},f}(i)$
     - ii. Otherwise it was assumed that $Q_{\mathrm{BF}}(i)=Q(i)$
2. If $Q(i)\geq Q(i-1)$ (i.e. rising limb):
   - (a) If $0<Q_{\mathrm{BF},b}(i)<Q_{\mathrm{BF},f}(i)$ then the backwards estimate was adopted: $Q_{\mathrm{BF}}(i)=Q_{\mathrm{BF},b}(i)$
   - (b) Otherwise the forward estimate was adopted: $Q_{\mathrm{BF}}(i)=Q_{\mathrm{BF},f}(i)$

For each of the station, a single best estimate of k_{BF} was derived from baseflow recessions following the method described in van Dijk (submitted, 2009). In that publication a non-linear filter was also tested and produced very similar results.

CC BY

Acknowledgement. This work is part of the water information research and development alliance between CSIRO's Water for a Healthy Country Flagship and the Bureau of Meteorology. The streamflow and catchment attribute data set used in this study was brought together by J. P. Guerschman, J. Peña Arancibia, Y. Liu and S. Marvanek of CSIRO Land and Water; their effort is gratefully acknowledged. This manuscript has benefited from reviews by P. Hairsine and Y. Zhang, also of CSIRO Land and Water.

References

Akaike, H.: Statistical predictor identification, Ann. Inst. Stat. Math., 22, 203–217 doi:10.1007/BF02506337, 1970.

Bergström, S.: The HBV model – its structure and applications, SMHI RH, 32, 1992

Beven, K. J.: Rainfall-Runoff Modelling: The Primer, John Wiley & Sons Inc, Chichester, UK, 372 pp. 2004.

Bloeschl, G.: Rainfall-runoff modelling of ungauged catchments, Article 133, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M. G., John Wiley & Sons, Chichester, UK, 2061–2080.2005.

Bonan, G. B.: A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: technical description and user's guide, National Center for Atmospheric Research, 1996

Chiew, F. H. S., Peel, M. C., and Western, A. W.: Application and testing of the simple rainfall-runoff model SIMHYD, in: Mathematical Models of Small Watershed Hydrology and Applications, edited by: Singh, V. P. and Frevert, D. K., Water resources Publication Littleton, Colorado, USA, 335–367, 2002.

Criss, R. E. and Winston, W. E.: Do Nash values have value? Discussion and alternate proposals, Hydrol. Process., 22, 2723–2725, 2008.

Guerschman, J.-P., van Dijk, A. I. J. M., McVicar, T. R., van Niel, T. G., Li, L., Liu, Y., and Peña Arancibia, J.: Water balance estimates from satellite observations over the Murray-Darling Basin, CSIRO, Canberra, Australia, 102 pp., http://www.clw.csiro.au/publications/waterforahealthycountry/mdbsy/, 2008.

Latron, J. and Gallart, F.: Runoff generation processes in a small Mediterranean research catchment (Vallcebre, Eastern Pyrenees), J. Hydrol., 358, 206–220, 2008.

Legates, D. R. and McCabe, G. J.: Evaluating the use of goodness of fit measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35, 233–241, 1999.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs, J. Geophys. Res., 99(D7), 14415–14428, 1994.

Liang, X. and Xie, Z.: A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models, Adv. Water Resour., 24, 1173–1193, 2001.

Liu, Y., van Dijk, A. I. J. M., Weerts, A. H., and De Jeu, R. A. M.: Comparison of Soil Moisture Simulated by HBV96 and Estimated from TRMM Passive Microwave Observations for a Catchment in Southern NSW, Australia, MODSIM 2007 International Congress on Modelling and Simulation. Christchurch, NZ, 2007, 1723–1728, 2007.

Maidment, D. R. E.: Handbook of Hydrology, McGraw-Hill, New York, USA, 1424 pp., 1992.

Mishra, S. K. and Singh, V. P.: Soil Conservation Service Curve Number (SCS-CN) Methodology, Water Science and Technology Library, 536 pp., 2003.

Oleson, K. W., Dai , Y., et al.: Technical description of the Community Land Model (CLM), 173, 2004.

Peel, M. C., Chiew, F. H. S., Western, A. W., and McMahon, T. A.: Extension of Unimpaired Monthly Streamflow Data and Regionalisation of Parameter Values to Estimate Streamflow in Ungauged Catchments. Report prepared for the Australian National Land and Water Resources Audit., Centre for Environmental Applied Hydrology. The University of Melbourne, 2000.

Peña Arancibia, J., van Dijk, A. I. J. M., Austin, J., and Stenson, M.: Predicting Changes in Streamflow and Salinity Patterns after Afforestation in the Southwest Goulburn region, Australia: An assessment of the water and salt export model 2CSalt, CSIRO Land and Water, Canberra, 2007.

Schoups, G., N. C. van de Giesen, and H. H. G. Savenije, Model complexity control for hydrologic prediction, Water Resour. Res., 44, W00B03, doi:10.1029/2008WR006836, 2008.

Surawski, N. and Yu, B.: Alternative probabilistic exponential distributions for modelling rainfall intensity In Australia, Proceedings, MODSIM International Congress on Modelling and Simulation, Melbourne, Australia, 2005.

Tanaka, T.: Storm runoff processes in a small forested drainage basin, Environ. Geol., 19, 179–191, 1992.

USDA-SCS: Section 4 – Hydrology, in: National Engineering Handbook, US Department of

Agriculture, Soil Conservation Service, Washington, D.C., 1985.
van Dijk, A. I. J. M. and Bruijnzeel, L. A.: Modelling rainfall interception by vegetation of variable density using an adapted analytical model. Part 2. Model validation for a tropical upland mixed cropping system, J. Hydrol., 247, 239–262, 2001.
van Dijk, A. I. J. M. and Bruijnzeel, L. A.: Runoff and soil loss from bench terraces. 1. An event-based model of rainfall infiltration and surface runoff, Eur. J. Soil Sci., 55(2), 299–316, 2004.
van Dijk, A. I. J. M., Bruijnzeel, L. A., Vertessy, R. A., and Ruijter, J.: Runoff and sediment generation on bench-terraced hillsides: measurements and up-scaling of a field-based model, Hydrol. Process., 19, 1667–1685, 2005a.
van Dijk, A. I. J. M., Meesters, A. G. C. A., Schellekens, J., and Bruijnzeel, L. A.: A two-parameter exponential rainfall depth-intensity distribution applied to runoff and erosion modelling, J. Hydrol., 300, 155–171, 2005b.
Wittenberg, H.: Baseflow recession and recharge as nonlinear storage processes, Hydrol. Process., 13, 715–726, 1999.
Yu, B., Rose, C. W., Coughlan, K. J., and Fentie, B.: Plot-scale rainfall-runoff characteristics and modeling at six sites in Australia and Southeast Asia, Trans. ASAE, 40, 1295–1303, 1997.

Table 1. Distribution of calculated (k_{QF}) and optimised values (others) values of the preferred model variants with three and four parameters, respectively. Note the units have been adjusted in two cases for ease of notation.

Parameter	n/a	C_2	C_4	C_5	C_6
Model variant	n/a	3 par.	3 par.	3 par.	4 par.
Symbol	k_{QF}	L_i	S_{max}	C_{Sg}	n/a
Unit	–	mm	10^3 mm	10^{-3} mm^{-1}	–
Median	0.532	11.6	5.67	6.48	0.91
Mean	0.524	22.9	145	12.0	0.942
CV	21%	149%	225%	2.86	60%
25–75% range	0.452–0.597	7.1–22.3	1.94–26.4	2.86–14.1	0.621–1.17
10–90% range	0.385–0.659	5.2–48.1	737–1000	1.01–34.6	0.362–1.54

HESSD

6, 5753–5782, 2009

Selection of an appropriately simple storm runoff model

A. I. J. M. van Dijk

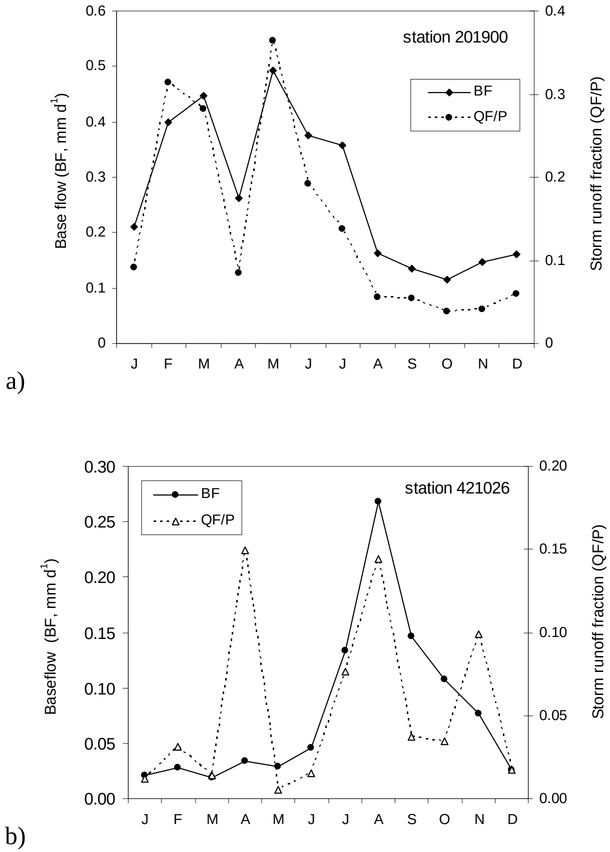

Fig. 1. Patterns in long-term (>5 years) average seasonal patterns in average baseflow (BF in $\mathrm{mm\,d^{-1}}$) and storm runoff fraction (QF/P): **(a)** station 201900 (Tweed River @ Uki on the NSW coast) experiencing $1475\,\mathrm{mm\,y^{-1}}$ with a somewhat drier spell from July–October, **(b)** station 421026 (Turon River @ Sofala in inland NSW) experiencing $763\,\mathrm{mm\,y^{-1}}$ rainfall with a dry spell from March–May. Note that there were no corresponding seasonal patterns in rainfall intensity.

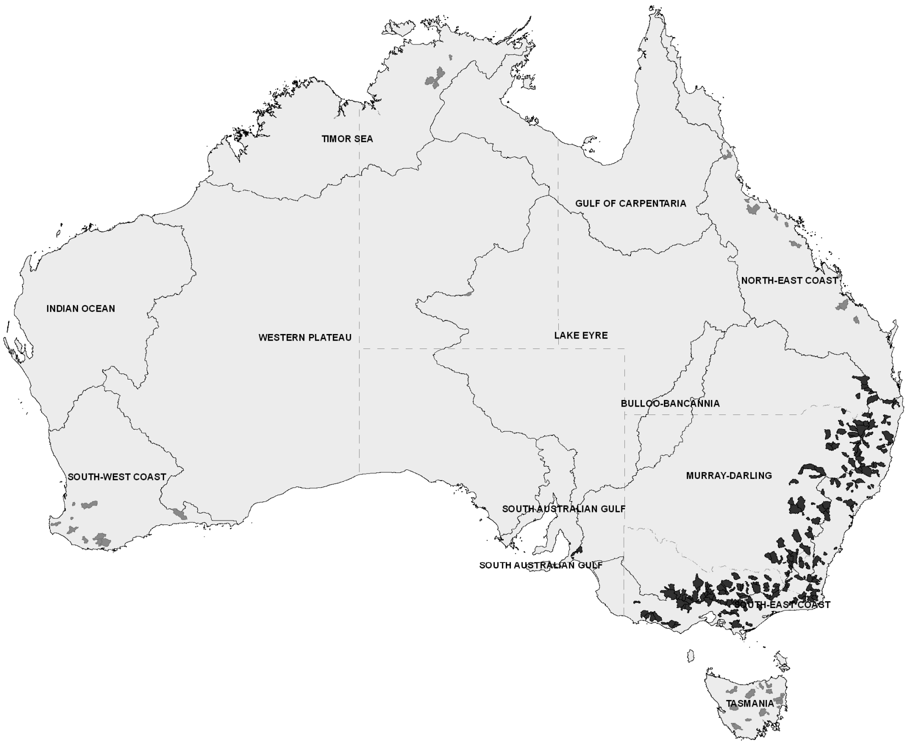

Fig. 2. Map showing the geographical distribution of the catchments for which streamflow data was used in this study.

HESSD

6, 5753–5782, 2009

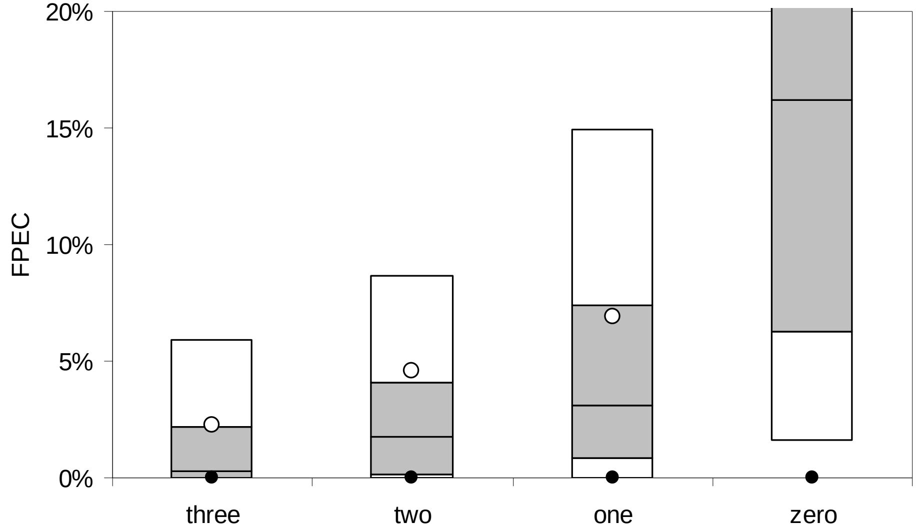

Fig. 3. Increase in the final prediction error criterion for each of the stations as the number of free parameters in the model structure is reduced.

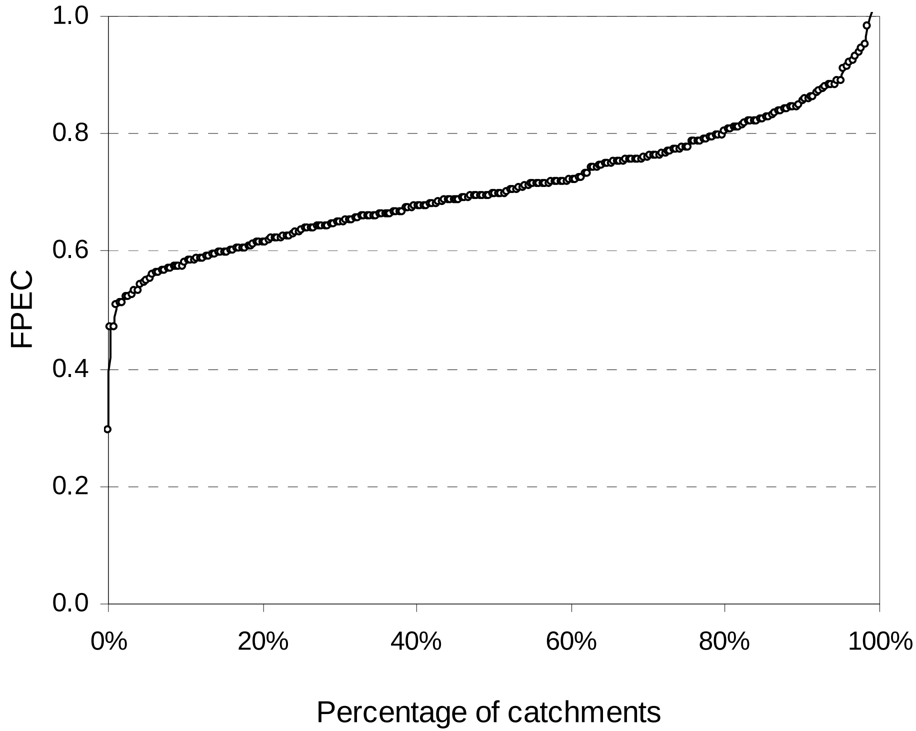

Fig. 4. Cumulative distribution of the final prediction error (FPEC) of the optimal model structure for all 260 catchments (FPEC is expressed here as the expected error sum of absolute differences between predicted and actual event storm runoff divided by total period storm runoff).

HESSD

6, 5753–5782, 2009

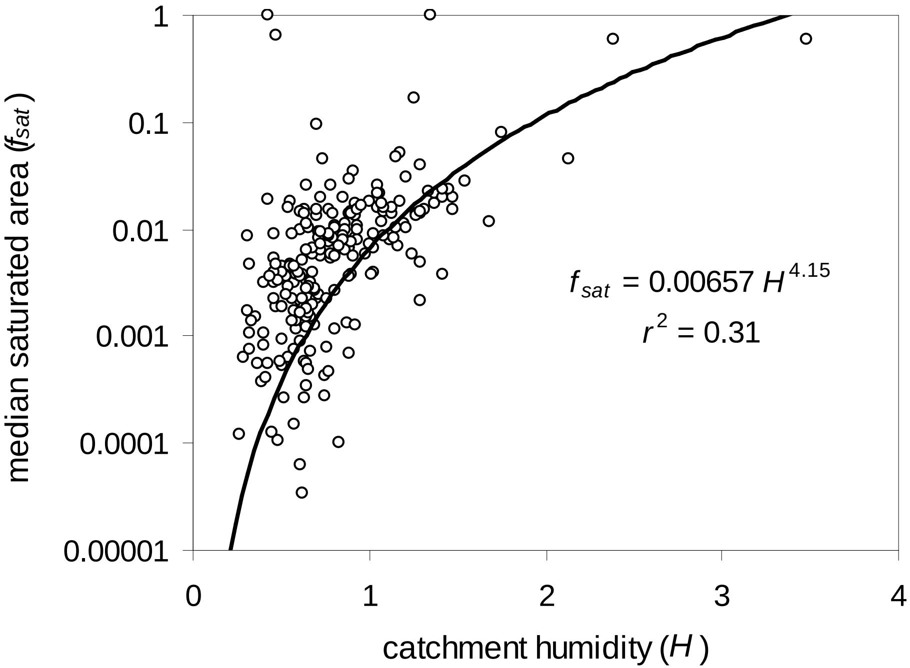

Fig. 5. Relationship between catchment humidity index (H, the ratio of rainfall over potential evaporation) and the median estimated fraction of saturated area (exceeded half of the time).

HESSD

6, 5753–5782, 2009

Selection of an appropriately simple storm runoff model

A. I. J. M. van Dijk