Climate Dynamics (1996) 12:685–699

Climate Dynamics

Testing a GCM land surface scheme against catchment-scale runoff data

A. J. Pitman[1], F. H. S. Chiew[2]

[1] School of Earth Sciences, Macquarie University, North Ryde, 2109, NSW, Australia
[2] Department of Civil & Environmental Engineering, University of Melbourne, Melbourne, Australia

Received: 9 June 1995 / Accepted: 4 April 1996

Abstract. A GCM land surface scheme was used, in off-line mode, to simulate the runoff, latent and sensible heat fluxes for two distinct Australian catchments using observed atmospheric forcing. The tropical Jardine River catchment is 2500 km^2 and has an annual rainfall of 1700 mm y^{-1} while the Canning River catchment is 540 km^2, has a Mediterranean climate (annual rainfall of 800 mm y^{-1}) and is ephemeral for half the year. It was found that the standard version of a land surface scheme developed for a GCM, and initialised as for incorporation into a GCM, simulated similar latent and sensible heat fluxes compared to a basin-scale hydrological model (MODHYDROLOG) which was calibrated for each catchment. However, the standard version of the land surface scheme grossly overestimated the observed peak runoff in the wet Jardine River catchment at the expense of runoff later in the season. Increasing the soil water storage permitted the land surface scheme to simulate observed runoff quite well, but led to a different simulation of latent and sensible heat compared to MODHYDROLOG. It is concluded that this 2-layer land surface scheme was unable to simulate both catchments realistically. The land surface scheme was then extended to a three-layer model. In terms of runoff, the resulting control simulations with soil depths chosen as for the GCM were better than the best simulations obtained with the two-layer model. The three-layer model simulated similar latent and sensible heat for both catchments compared to MODHYDROLOG. Unfortunately, for the ephemeral Canning River catchment, the land surface scheme was unable to time the observed runoff peak correctly. A tentative conclusion would be that this GCM land surface scheme may be able to simulate the present day state of some larger and wetter catchments but not catchments with peaky hydrographs and zero flows for part of the year. This conclusion requires examination with a range of GCM land surface schemes against a range of catchments.

Correspondence to: A. J. Pitman

1 Introduction

The importance of land surface processes in climate modelling has long been recognised (e.g. Mintz 1984).The parametrisation of hydrological processes in climate models, and in particular general circulation models (GCMs), has been the subject of a vast amount of research over the last decade following pioneering work by, for instance, Manabe (1969) and Deardorff (1977, 1978). More recently Dickinson et al. (1986, 1993), Sellers et al. (1986), Noilhan and Planton (1989), Abramopoulos et al. (1988) and many other groups have developed land surface parametrisation schemes for inclusion into GCMs. The importance of the parametrisation of the land surface was illustrated in Gates et al. (1990) and the role of runoff parametrisation in affecting GCM simulations was discussed by Viterbo and Illari (1994). In addition, Nobre et al. (1991) and Henderson-Sellers et al. (1993) examined the role of the land surface in deforestation experiments, Xue and Shukla (1993) in desertification experiments and Whetton et al. (1994) in experiments which attempted to predict how runoff characteristics might change in a greenhouse-warmed world. In developing land surface schemes, attempts have been made to validate these models by testing them against meteorological data (e.g. Sellers et al. 1989; Noilhan and Planton 1989). Attempts have also been made to compare land surface models against each other to identify outliers or unusual behaviour (e.g. Pitman et al. 1993b; Polcher et al. 1995; Bonan 1994).

While the fluxes of latent and sensible heat are important quantities simulated by GCM land surface models, these quantities are difficult to measure over time scales appropriate to climate change (years to decades). Therefore, land surface modellers have looked for alternative sources of data with which to validate and improve land surface models. Arnell (1995), for instance, reviews the value of runoff data as a validation tool for atmospheric models. Rowntree and Lean (1994) used runoff data to examine the model developed by Warrilow et al. (1986) against runoff observations from southeast England and conclude that the amount of runoff simulated was deficient due to the point-based nature of the model and the consequent lack of heterogeneity in the runoff generation processes. They also found that an alternative model which did include heterogeneity, if calibrated carefully, could generally reproduce the observed runoff.

In a related paper (Chiew et al. 1996), simulations from a conceptual rainfall-runoff model (MODHYDROLOG) and a general circulation model (GCM) land surface scheme (Bare Essentials of Surface Transfer, BEST, Cogley et al. 1990; Pitman et al. 1991) were compared. A brief summary of MODHYDROLOG is provided in Appendix A and Fig. 1 illustrates the conceptual design of both models. In this work, a series of more detailed experiments are performed to improve the runoff simulated by BEST for the Jardine River and Canning River catchments (see Fig. 2). The Jardine River catchment at Telegraph Line (11°09′S, 152°10′E) is located in northern Queensland and has a drainage area of 2500 km^2. It has a tropical climate with an annual rainfall of 1700 mm, more than 90% of which occurs in summer and autumn (December to May). More than half of its annual rainfall becomes runoff. The Canning River catchment (32°14′S, 116°10′E) is located in southwest Western Australia and drains an area of 540 km^2. It has a Mediterranean climate and is dominated by winter rainfall with more than 50% of its annual rainfall of 800 mm occurring in winter (June to August). Its runoff coefficient is low with annual runoff being less than 3% of rainfall. The catchment is ephemeral with the monthly streamflow volume being zero for half of the year. A more detailed description of both catchments is given in Chiew et al. (1996).

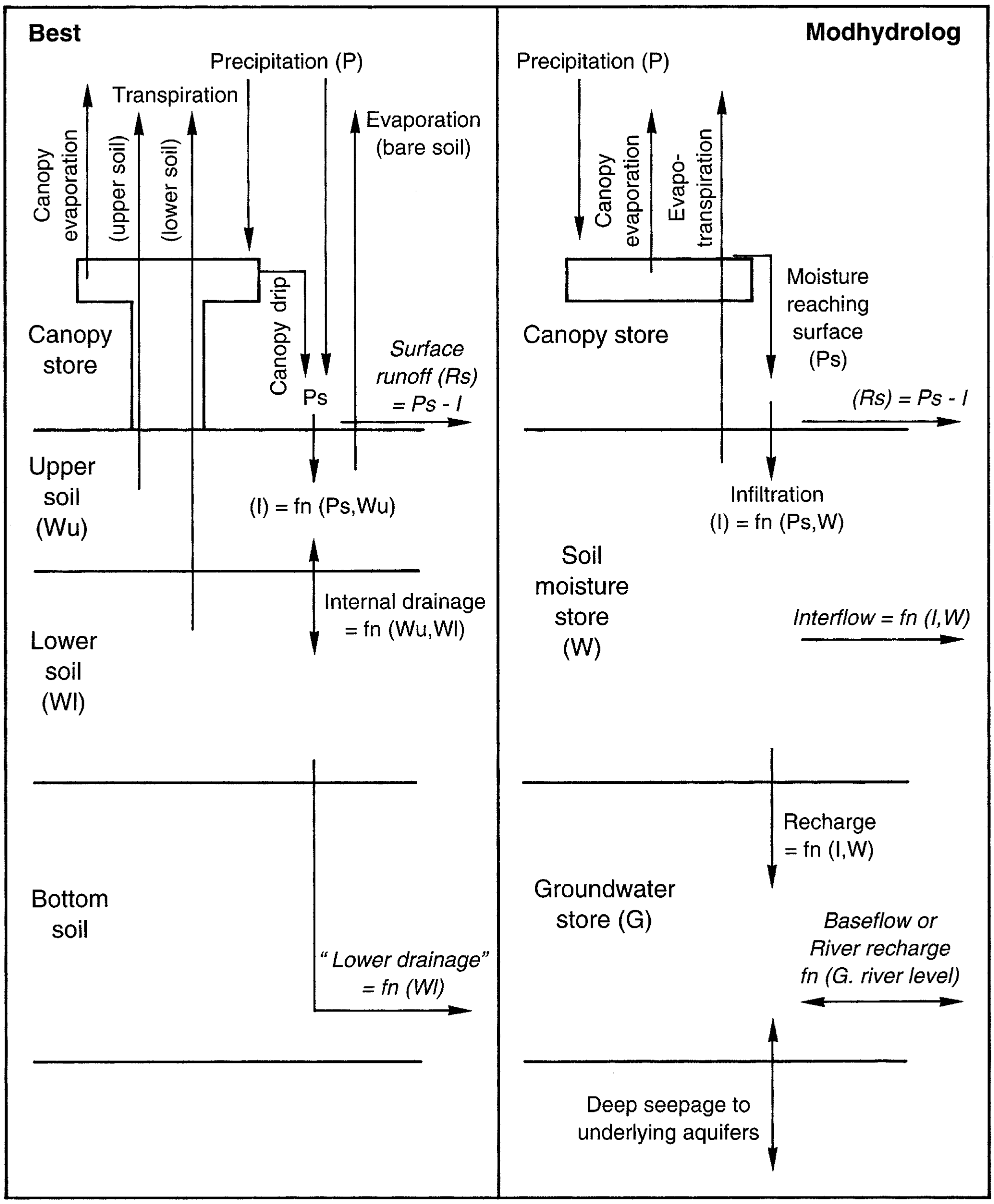

Fig. 1. Representation of soil water and runoff by BEST and MODHYDROLOG. W refers to soil moisture stores, *arrows* show fluxes, I refers to infiltration and P to precipitation. Runoff terms are shown in italics. W (W_u and W_l) and groundwater store (G) are ratios of moisture levels to capacity or saturation

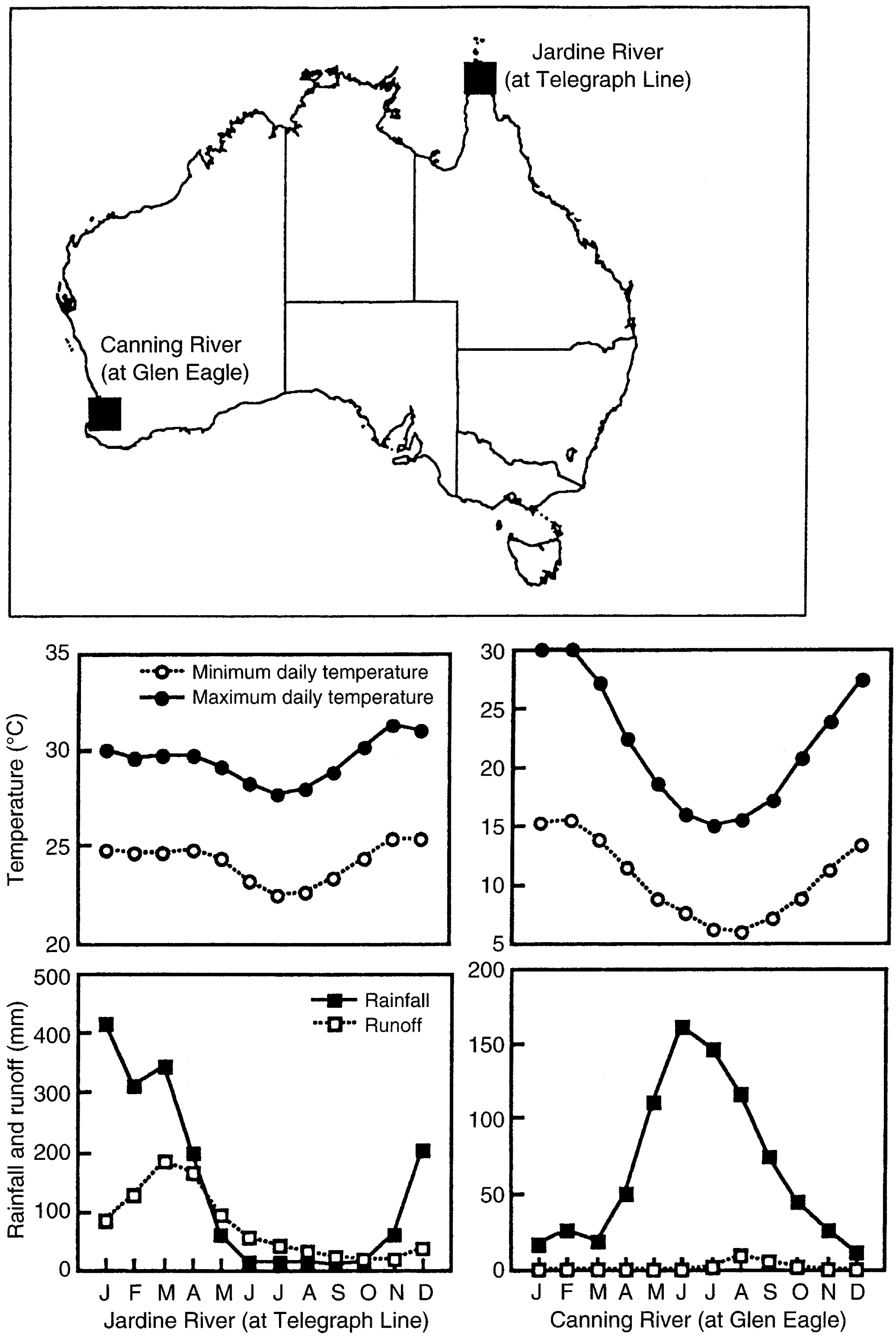

Fig. 2. Geographical location and selected climate data for the two locations used in these experiments

2 Experimental design

The simulations were performed using observed climate forcing at 30 min intervals. This forcing included precipitation, solar radiation, infrared radiation, air temperature, wind speed, surface pressure and vapour pressure which BEST used to simulate surface temperature, soil moisture, runoff and turbulent energy fluxes. The rainfall data were obtained directly from 30-min pluviograph records. The surface pressure, air temperature, specific humidity and wind speeds were interpolated from observations taken every three hours to the 30 min resolution required. Half hourly solar radiation was available from the Australian Bureau of Meteorology. Finally the incoming longwave radiation was estimated from surface air temperature, specific humidity and cloud cover using the Brunt equation (Budyko 1974). These simulations were for

11 years for the Canning River (gauging station number 616065) and 16 years for the Jardine River (gauging station number 927001). The first year from all simulations were omitted from subsequent analysis to avoid initialisation problems.

This method of assessing land surface models prevents any feedbacks between the surface and the atmosphere in that, whatever happens to the land surface, the atmospheric forcing is not affected. The results and sensitivities described in this study are therefore only the first order results and cannot be extrapolated to the fully coupled (GCM) environment. Despite obvious problems with this method of using a land surface model (see Dolman and Gregory 1992; Jacobs and de Bruin 1992; Pitman et al. 1993a) 'standalone' testing of models is widely used in the climate modelling community (Sellers and Dorman 1987; Abramopoulos et al. 1988; Sellers et al. 1989; Verseghy 1991; Mihailovic et al. 1992; Pitman et al. 1993b; Verseghy et al. 1993) and given that it is not feasible to couple MODHYDROLOG to a GCM we use this methodology here, but interpret the results and their implications with care.

The first set of experiments described in this study were performed with the basic version of BEST which utilises a two-layer hydrological model, as coupled into the Bureau of Meteorology Research Centre (BMRC) GCM (Yang et al. 1995) and using parameter data appropriate to the GCM grid square containing the specific catchment. This version of BEST performs comparably with other land surface models collaborating in the Project for the Intercomparison of Land Surface Parametrisation Schemes (PILPS, e.g. Pitman et al. 1993b). A second set of experiments were then performed in order to try and improve the predictive skill of BEST by calibrating specific parameters for the two catchments. The aim in these simulations was to simulate the observed runoff as well as possible since this was the only observed quantity available for validation. In addition, the latent and sensible heat fluxes estimated by MODHYDROLOG are compared with those simulated by BEST although as discussed by Chiew et al. (1996) it is difficult to make reliable conclusions about the monthly and seasonal estimates of the turbulent energy fluxes.

In this study a series of experiments are therefore described. For a two-layer version of BEST, a control experiment is conducted, followed by a series of sensitivity experiments. These will be discussed first, followed by experiments with the extended version of BEST.

3 Experiments with a two-layer model

3.1 Simulating the Jardine River catchment runoff using a 2-layer model

The control simulation from BEST (prior to calibration), the simulation by MODHYDROLOG and the final results from BEST (after calibration) are shown in Fig. 3. Since the albedo has been prescribed differences between the latent and sensible heat fluxes simulated by BEST or MODHYDROLOG are not due to errors in the simulation of net radiation. Figure 3a shows runoff, and as expected, MODHYDROLOG simulates the observed runoff well (it is calibrated to do so for this specific catchment). In contrast the control version of BEST, using the default parameters appropriate to a GCM grid square containing this catchment, overestimates runoff early in the year (during periods high rainfall in January, February and March) leading to an underestimation of runoff later in the year. Figure 3b, c shows that this poor simulation of runoff impacts relatively little on the simulated turbulent energy fluxes which are similar to MODHYDROLOG except when MODHYDROLOG simulates sustained negative sensible heat fluxes.

A series of experiments were then conducted in order to calibrate BEST and reduce the runoff peak in February and March, and to increase the gravitational drainage in later months. The poor simulation of runoff was believed due to an underestimation of the soil water storage capacity leading to an inability to store enough water during heavy rain, resulting in insufficient soil water to support gravitational drainage during periods of low rainfall. Figure 4a shows the control version of BEST where the upper soil depth was set at 0.1 m and the lower depth at 1.9 m (this is the same curve as shown in Fig. 3a). Increasing the upper soil layer to 0.5 m and the lower layer to 2.5 m reduced the runoff peak (Fig. 4b) although the lack of runoff between July and November is not supported by the observations. As a consequence of the changes in soil depths, the latent heat (Fig. 5b) and sensible heat fluxes simulated by BEST and MODHYDROLOG were quite different. The deeper upper soil layer in BEST retained more moisture and reduced the high runoff peak but this led to a higher latent heat flux throughout the year and a lower sensible heat flux (not shown) since the latent heat flux did not become moisture limited.

In an attempt to reduce the runoff peak but retain the lower runoff amounts later in the year, the upper soil layer depth was increased to 0.8 m, the lower layer was increased to 8.0 m and the hydraulic conductivity was also increased to 0.0065 mm s^{-1} (typical of loam, Dickinson et al. 1993). The result (Fig. 4c) was an underestimated runoff peak but the simulation of some runoff throughout the year in agreement with the observations. The magnitude of the runoff peak could be increased (Fig. 4d) by reducing the depth of the lower soil layer to 5.0 m but this did not impact significantly on the latent heat flux (Fig. 5d). Using a lower soil depth of 8.0 m and a hydraulic conductivity of 0.1 mm s^{-1} (typical of sand, Dickinson et al. 1993) improved the runoff simulation further (Fig. 4e) but the latent heat flux simulated by BEST (Fig. 5e) was still very different to that simulated by MODHYDROLOG.

From Figs. 4 and 5 it was concluded that no reasonable combination of soil depths and hydraulic conduc-

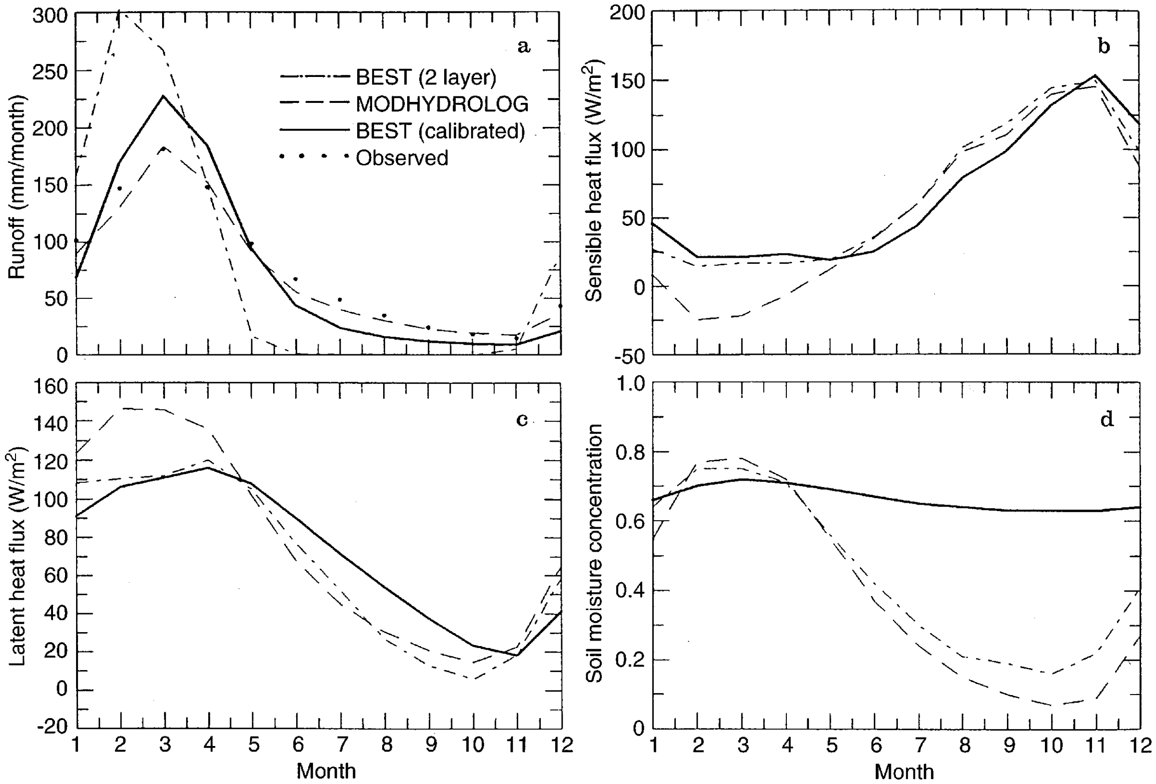

Fig. 3a–d. Average annual cycle (monthly averages over length of simulation) for the Jardine River catchment of **a** runoff by the two-layer default version of BEST (*dash-dot line*) and MODHYDROLOG (*dashed line*). The *dots* show observed runoff (all in mm month^{-1}). The simulation by BEST with parameters calibrated to reproduce the observed runoff is shown by the *solid line,* **b** as **a** except for sensible heat flux, W m^{-2}; **c** latent heat flux, W m^{-2}; and **d** soil moisture concentration (ratio of saturation)

tivity was able to reproduce **both** the observed runoff and the latent heat flux simulated by MODHYDROLOG. BEST required a thin upper soil layer in order to simulate a similar seasonality in the latent heat flux when compared to MODHYDROLOG but also required a thick layer to store enough water to simulate the observed runoff. In the final simulation therefore, the upper soil layer depth was returned to 0.1 m, while retaining the 8.0 m lower soil layer in an attempt to provide higher water storage. The result (Fig. 4f) was a reasonable simulation of runoff and a simulation of the latent heat flux (Fig. 5f) somewhat closer to MODHYDROLOG. Attempts to improve this simulation by modifying the root distribution, fractional vegetation cover and stomatal resistance formulation were unsuccessful.

Figure 3 shows all the quantities from the final simulation and compares them with the original control simulation. The runoff simulated by BEST with soil depths of 0.1 m and 8 m is reasonable although the peak in March is overestimated by 50 mm month^{-1}. The seasonal shape and variability in the turbulent energy fluxes are similar to MODHYDROLOG, but at the expense of prescribing a very deep lower soil depth of 8.0 m (which is the maximum depth noted by Webb et al. 1993 but is smaller than the 10 m depth used by BATS, Dickinson et al. 1993). The combination of the 0.1 m upper soil layer and 8.0 m lower depth prevents any seasonality in soil moisture levels (Fig. 3d) which is present in both the control version of BEST and the simulations by MODHYDROLOG. This simulation of soil moisture, sensible heat, latent heat and runoff are also very different to MODHYDROLOG although in the case of soil moisture this is partly due to the different soil depths and water holding capacity used in the two models. These differences should be considered in light of the results of Shao et al. (1994) who show that there appears to be little correlation between the simulation of soil moisture and latent heat in a number of land surface models (because the functional relationship between the latent heat flux and soil moisture is the important factor, not the actual soil moisture).

3.2 Simulating the Canning River catchment runoff using a 2-layer model

The simulations for the ephemeral Canning River catchment are rather more demanding than the Jardine River catchment since the land surface has to respond to low frequency but quite high intensity rainfall events. The control version of BEST (where no calibration has taken place) simulates the observed runoff badly (Fig. 6a) as a runoff peak of more than 50 mm

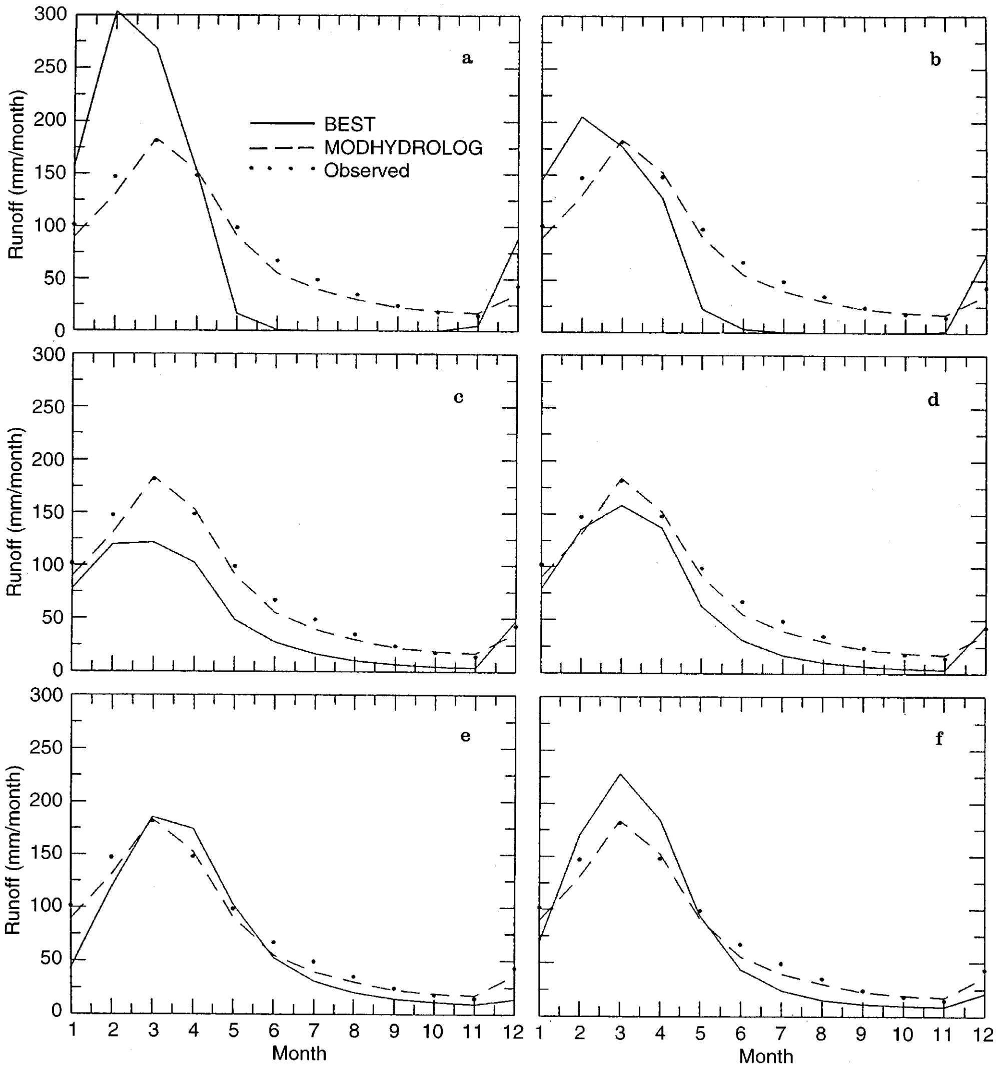

Fig. 4a–f. As Fig. 3a but with various soil depths and hydraulic conductivities. In all cases the simulation by MODHYDROLOG is shown as a *dashed line* and the observed data are shown as *dots*. The parameter values are: **a** $d_u=0.1$ m, $d_l=1.9$ m, $K_H=0.005$ mm s^{-1}; **b** $d_u=0.5$ m, $d_l=2.5$ m, $K_H=0.005$ mm s^{-1}; **c** $d_u=0.8$ m, $d_l=8.0$ m, $K_H=0.0065$ mm s^{-1}; **d** $d_u=0.8$ m, $d_l=5.0$ m, $K_H=0.0065$ mm s^{-1}; **e** $d_u=0.8$ m, $d_l=8.0$ m, $K_H=0.1$ mm s^{-1}; **f** $d_u=0.1$ m, $d_l=8.0$ m, $K_H=0.1$ mm s^{-1}; where d_u is the upper soil depth, d_l is the lower soil depth and K_H is the saturated hydraulic conductivity. Note that the *curves* shown in **a** and **f** are reproduced in Fig. 3a

month^{-1} is simulated, in contrast to an observed runoff rate of less than 10 mm month^{-1}. The simulation by BEST of sensible (Fig. 6b) and latent heat (Fig. 6c) are also very different to those simulated by MODHYDROLOG.

A significant improvement was obtained by increasing the upper soil depth to 0.6 m and the lower soil depth to 3.0 m (Fig. 7b). Runoff was reduced to the observed magnitude, but the main peak occurred too late. In order to simulate the runoff peak at the right time, the soil depths were increased to 1.2 m (upper) and 3.5 m (lower) resulting in runoff occurring at the right time, but being under predicted (Fig. 7c). Lastly, the depths were reduced to 1.1 m (upper) and 3.2 m (lower) which marginally improved the runoff simulation (Fig. 7d) although the latent heat flux simulation (Fig. 8d) remained very different from MODHYDROLOG.

The remaining variables for this final simulation are all shown in Fig. 6. BEST consistently predicts higher sensible and lower latent heat fluxes during the rainfall season compared to MODHYDROLOG (Fig. 6c). MODHYDROLOG simulated negative sensible heat during the rainfall season (warming the surface) which BEST did not simulate. This negative sensible heat permitted MODHYDROLOG to simulate a much higher latent heat flux (in excess of net radiation) which is not realistic (see Chiew et al. 1996).

The two-layer version of BEST is clearly unable to simulate the combination of runoff, soil moisture and the turbulent energy fluxes realistically. In order to simulate the turbulent energy fluxes, of primary importance for the GCM, BEST requires a thin layer in contact with the atmosphere (cf. Mahrt and Pan 1984). In contrast, BEST can simulate runoff realistically (of primary importance to hydrologists) if the soil depths are calibrated. One solution to this problem is to extend BEST to three layers for soil hydrology.

4 Extension of BEST to a three soil layer hydrological model

This section describes the extension of BEST to a three soil layer hydrological model. Land surface parametrisation schemes often include more than two

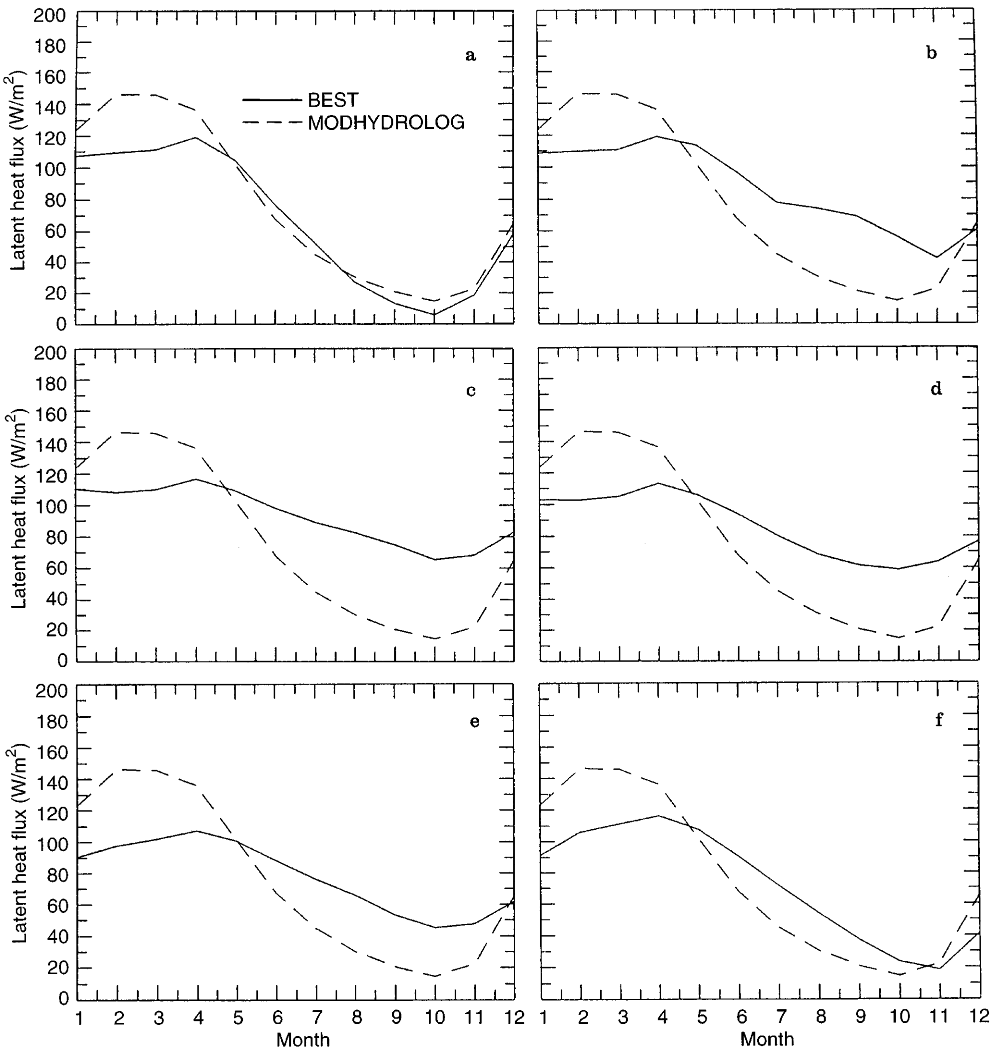

Fig. 5a–f. As for Fig. 4 but for the latent heat flux (W m^{-2})

layers (see Pitman et al. 1993b, Table 1) and in BEST this extension means that the same number of layers are used for both heat and moisture. The soil temperature model and the canopy parametrisation remain unchanged in these experiments.

The basic two layer soil moisture model is described in Appendix B of Chiew et al. (1996). In the three layer model the equations for the liquid soil moisture in the upper soil layer (W_U) and lower soil layer (W_L) remain unchanged while the soil moisture content of the bottom soil layer (W_B) becomes:

$$X_v \frac{dW_B}{dt} = \frac{R_{lb} - R_{bg}}{{}_w d_b} \qquad (1)$$

where X_v is the soil porosity, R_{bg} is the runoff via gravitational drainage to ground water, d_b is the depth of the bottom soil layer and $_w$ is the density of water. R_{lb} is the flux of water between the lower and bottom soil layers which is based on Darcy's equation for one-dimensional fluid flow and is defined as:

$$R_{lb} = K_{Hl}\left[1 - \left(\frac{\Delta\psi}{\Delta z}\right)_l\right] \qquad (2)$$

where z is depth and K_{Hl} is the hydraulic conductivity in the lower soil layer, calculated following Clapp and Hornberger (1978). The moisture potential (ψ) is:

$$\left(\frac{\Delta\psi}{\Delta z}\right)_l = B\,\psi_0\left(\frac{W_L + W_B}{2}\right)^{-B-1}\frac{W_L - W_B}{d_l} \qquad (3)$$

where ψ_0 is the soil water suction at saturation, B is the Clapp and Hornberger parameter and d_l is the depth of the lower soil layer. The capillary rise of soil water R_{bg} is written according to Darcy's law:

$$R_{bg} = K_{Hb}\left[1 - \left(\frac{\Delta\psi}{\Delta z}\right)_b\right] \qquad (4)$$

The hydraulic conductivity in the soil is given by

$$K_H = K_{H0}\,W^{2B+3} \qquad (5)$$

where K_H equals either K_{Hl} (if W equals W_L) or (K_{Hb} (if W equals W_B). The moisture potential gradient required in Eq. (4) is parametrized as:

$$\left(\frac{\Delta\psi}{\Delta z}\right)_b = \begin{cases} B\,\psi_0\,W_B^{-B-1}\left(\dfrac{W_B - W_{FC}}{d_b}\right) & W_B > W_{FC} \\ 0 & W_B \leq W_{FC} \end{cases} \qquad (6)$$

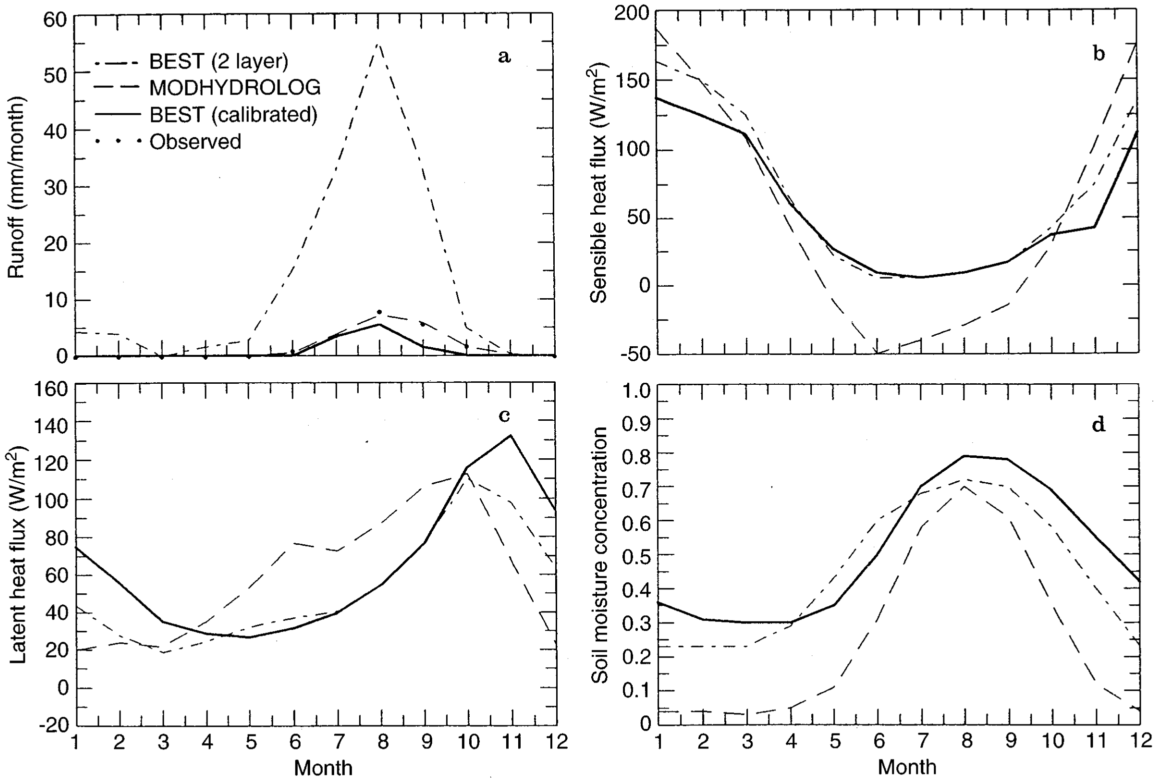

Fig. 6a–d. As for Fig. 3 but for the Canning River catchment

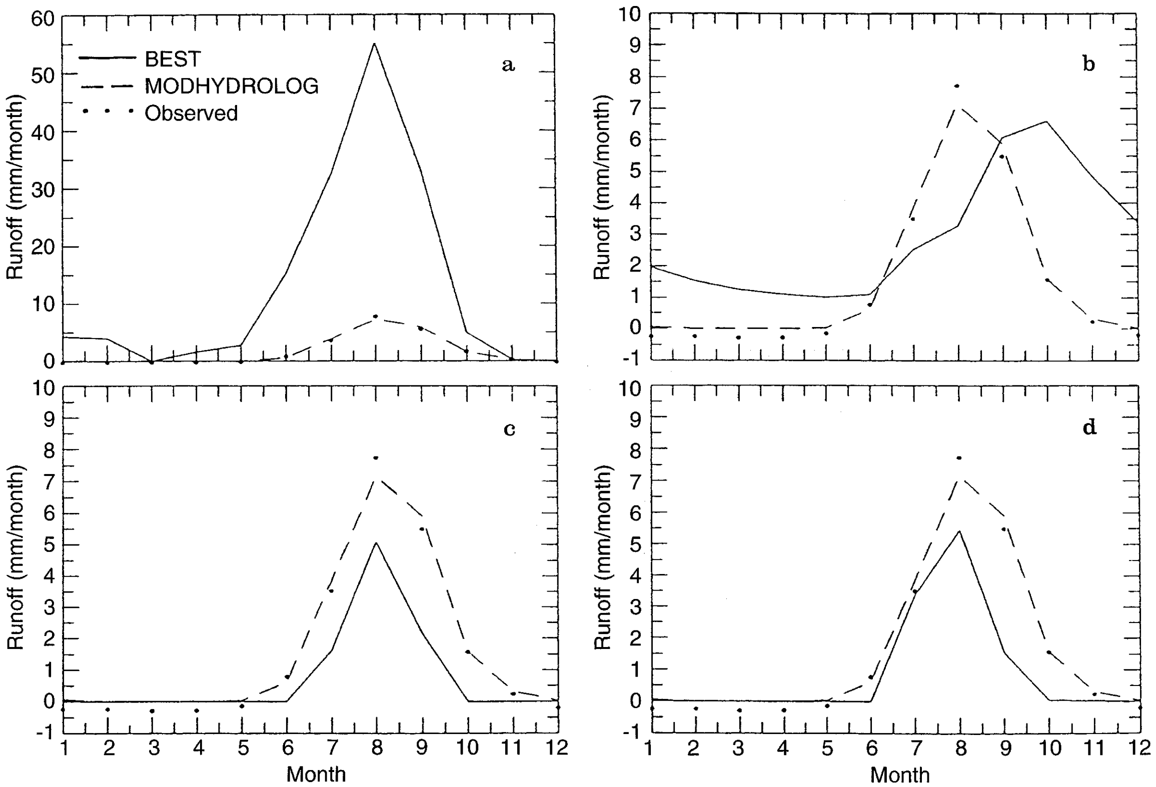

Fig. 7a–d. As Fig. 6a but with various soil depths and hydraulic conductivities. In all cases the simulation by MODHYDROLOG is shown as a *dashed line* and the observed data are shown as dots. The parameter values are: **a** $d_u=0.1$ m, $d_l=1.9$ m; **b** $d_u=0.6$ m, $d_l=3.0$ m; **c** $d_u=1.2$ m, $d_l=3.5$; **d** $d_u=1.1$ m, $d_l=3.2$ m where d_u is the upper soil depth and d_l is the lower soil depth. Note that the *Y-axis scale* of panel **a** differs from the other panels

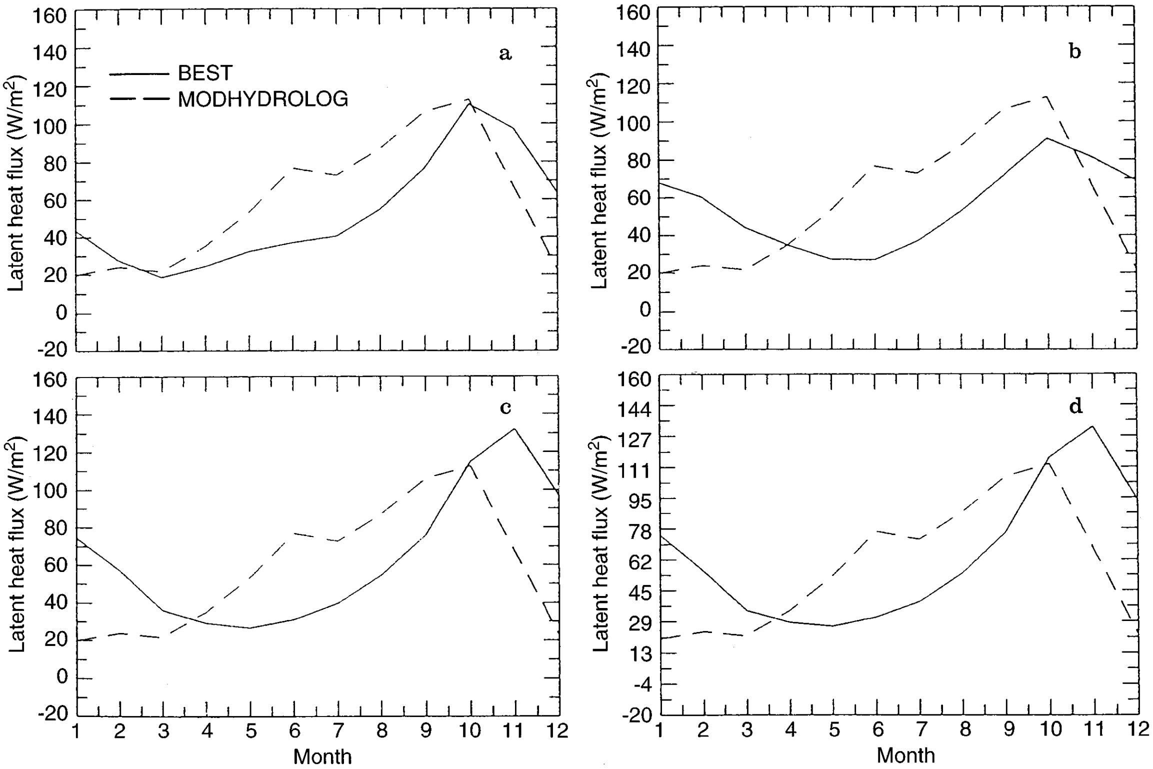

Fig. 8a–d. As for Fig. 7 but for the latent heat flux (W m^{-2})

where the moisture content in the layer beneath that containing W_B is denoted W_{FC} and is held at 'field capacity'. The total runoff (R_{tb}) simulated by BEST becomes

$$R_{tb} = R_{su} + R_{bg} \quad (7)$$

where R_{su} is the surface runoff (Eq. 12 of Chiew et al. 1996) and R_{bg} is the loss of water through gravitational drainage (Eq. 4).

In choosing the three layers, it was decided to retain the 0.1 m top layer, and 1.9 m second layer, and include the third layer as a deep layer of 5 m. A series of sensitivity tests showed that the model was not sensitive to the depth of the third layer in these catchments.

5 Experiments with a three-layer model

The experiments described in Sect. 3 were repeated with the three layer model.

5.1 Simulating the Jardine River catchment runoff using a 3-layer model

The extension of BEST to a three soil layer hydrological model led to an improved simulation of the Jardine River catchment. The control simulation (Fig. 9) can be compared with the control simulation using the two-layer model (Fig. 3) to show that the default simulation of runoff is greatly improved. The control runoff simulation with the three soil layer model reproduces the gradual reduction in runoff without changing soil depths or the hydraulic conductivity from the default values. However, the high runoff regime in April and May is slightly overestimated compared to the observations and an improved simulation was possible by decreasing the middle soil layer depth from 1.9 m to 1.7 m producing the final simulation for this catchment (Fig. 9). The simulation of the sensible heat (Fig. 9b) and the latent heat fluxes (Fig. 9c) both closely match those of MODHYDROLOG except when MODHYDROLOG simulated negative sensible heat.

The full 16 year simulation by BEST shows that runoff is well simulated including the peak discharge and the subsequent drying (Fig. 10a). BEST and MODHYDROLOG show a consistent difference in the simulation of monthly sensible heat (Fig. 10b) and latent heat (Fig. 10c) fluxes with MODHYDROLOG producing negative sensible heat every month which is not reproduced by BEST. As a result, MODHYDROLOG simulates a higher latent heat flux, driven by a combination of net radiation and negative sensible heat indicating that while the model simulates the annual runoff well, the method used by MODHYDROLOG for calculating latent and sensible heat is sometimes deficient.

Overall, BEST's simulation of this catchment was excellent which gives us confidence that, for this type of catchment and climatology, a three soil layer model will simulate runoff realistically.

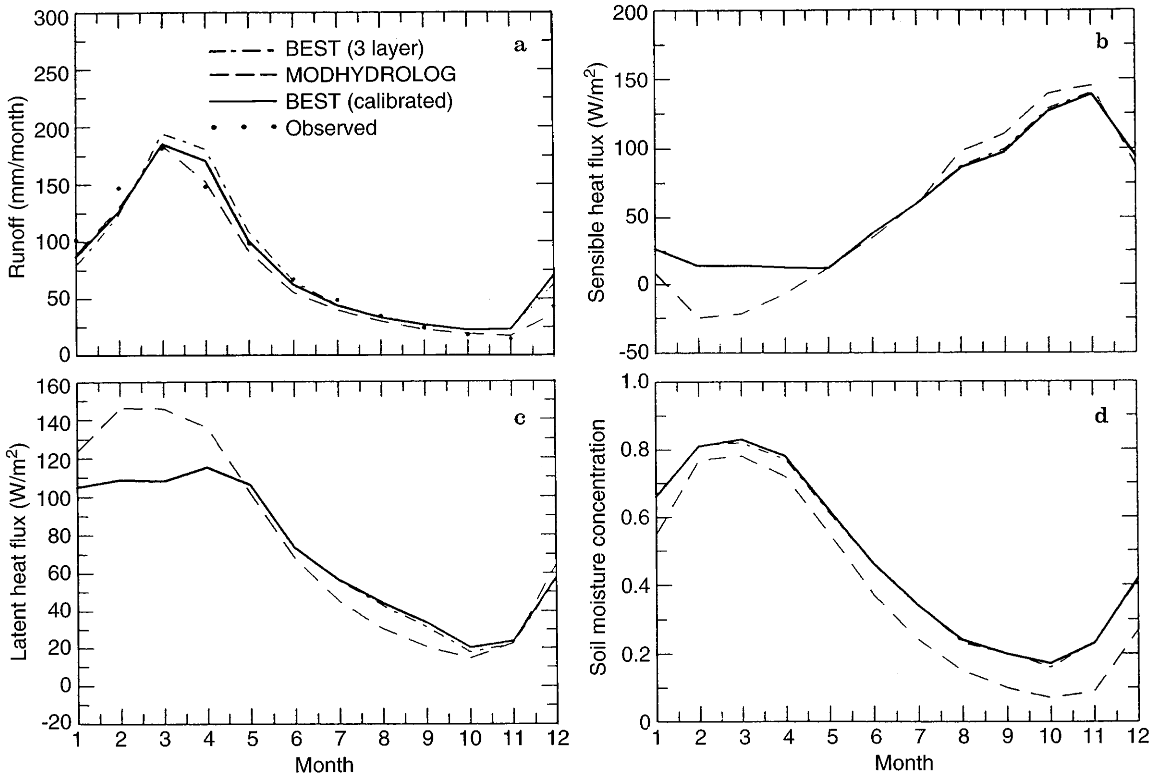

Fig. 9a–d. As for Fig. 3 but for the three layer version of BEST. The parameter values are $d_u = 0.1$ m for both runs with BEST and $d_l = 1.9$ m in the control and 1.7 in the final simulation. $K_H = 0.005$ mm s^{-1} in both simulations

5.2 Canning River catchment, 3-layer model

In contrast to the Jardine River catchment, the control simulation by the three-layer version of BEST of the Canning River catchment was unsuccessful. Figure 11 shows that while the control simulation by the three-layer version of BEST is superior to the control using a two-layer model (Fig. 6) BEST simulates runoff too early in the year, and simulates too much runoff overall compared to the observed.

Many experiments were conducted using this version of BEST, but it proved impossible to simulate the observed runoff successfully. Using an upper soil layer depth of 0.3 m produced Fig. 11a where the magnitude of the simulated runoff was good, but BEST persisted in simulating the runoff two months early because the existing algorithms in BEST cannot simulate the time delay in runoff (baseflow) adequately in ephemeral catchments with peaky hydrographs and zero flows for most of the year. Note however that the turbulent energy fluxes are generally quite similar to those simulated by MODHYDROLOG.

6 Discussion

The control simulation of runoff by the two-layer version of BEST were rather poor. Increasing the lower soil depths to 3.2 m in the case of the Canning River catchment and 8.0 m in the Jardine River catchment led to reasonable simulations of the average runoff. However, the turbulent energy fluxes remained quite different from those simulated by MODHYDROLOG due to the depths required to simulate the observed runoff realistically. If the control soil depths were chosen, similar turbulent energy fluxes were simulated to MODHYDROLOG, but BEST then failed to simulate the observed runoff reasonably. Therefore, with the two-layer model, either the observed runoff or the turbulent energy fluxes simulated by MODHYDROLOG could be simulated realistically.

The extension of BEST to a three-layer model improved the simulation of runoff. The control simulations with BEST were close to the observed (in the case of runoff) and close to the simulations by MODHYDROLOG in the case of turbulent energy fluxes. In the case of the Jardine River catchment, the control runoff simulation with the three-layer model was superior to the best simulation attained using the calibrated two layer model. The ability to retain the shallow top soil layer (0.1 m) also permitted a simulation of turbulent energy fluxes similar to MODHYDROLOG. Considering that Figs. 9 and 10 shows the simulation by BEST of a single catchment with no knowledge of the sub-surface catchment characteristics, it is noteworthy that BEST can simulate runoff well for this type of catchment.

BEST showed negligible skill in simulating runoff for the Canning River catchment. Figure 11a suggests that the problem in simulating the observed runoff for

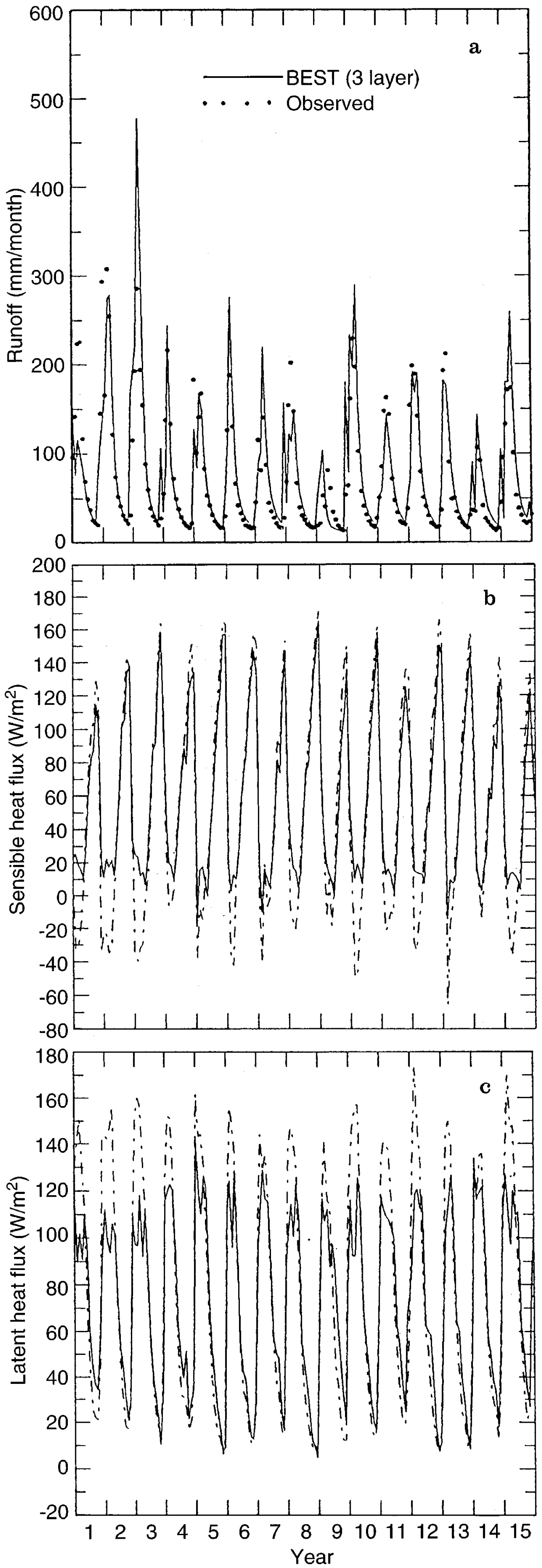

Fig. 10a–c. Simulation over the full 15-year record from BEST for the Jardine River Catchment of **a** runoff (mm month^{-1}); **b** sensible heat flux (W m^{-2}); and **c** the latent heat flux (W m^{-2}). Panel **a** shows BEST and the observed values while panels **b** and **c** show BEST and MODHYDROLOG

this catchment is because simulated runoff is not time lagged and appears at the edge of the "catchment" instantaneously. This simplification is common in GCM land surface schemes, although river routing and runoff lags are being introduced into some hydrological schemes intended for GCMs (e.g. Dümenil and Todini 1992). While BEST is able to simulate runoff resulting from all storms for the Jardine River catchment reasonably (Fig. 10), it cannot simulate any of the runoff peaks in the Canning River catchment. This may be related to basin characteristics, the distribution of rainfall type (i.e. convective or frontal), the 'spotty' nature of the rainfall, or the size of the Canning River catchment (540 km^2) which is small compared to the Jardine River (2500 km^2) and implies that microscale heterogeneities may influence the results proportionally more in a smaller catchment.

Rowntree and Lean (1994) also used runoff data to validate a land surface model and argued that the land surface scheme developed by Warrilow et al. (1986), which did not include the types of sub-grid scale catchment characteristics used by Dümenil and Todini (1992) was unable to simulate runoff for two catchments in southeast England. The Warrilow et al. (1986) model included a single soil layer for the calculation of runoff and, according to the results shown here, would not be able to simulate runoff realistically because this one (relatively deep) layer does not provide the vertical resolution required to simulate the important runoff fluxes. Rowntree and Lean (1994) argued the need to include a Dümenil and Todini (1992) type model but the results discussed here do not support this finding for wet catchments with high rainfall where runoff generation is basically a continuous process. However, without incorporating an additional term into the runoff generation processes, BEST proved unable to simulate the ephemeral Canning River catchment. While sub-grid scale runoff generation might represent this additional term, the main problem appears to be one of timing, in that rainfall occurs and runoff is immediately produced by the model. In the case of the Canning River catchment, there is about a month of significant rainfall before runoff is measured, during which moisture stores are filled. After this, infrequent but large rainfall events generate runoff, which is not simulated by BEST because of the lack of heterogeneity, a finding in support of the conclusions of Rowntree and Lean (1994).

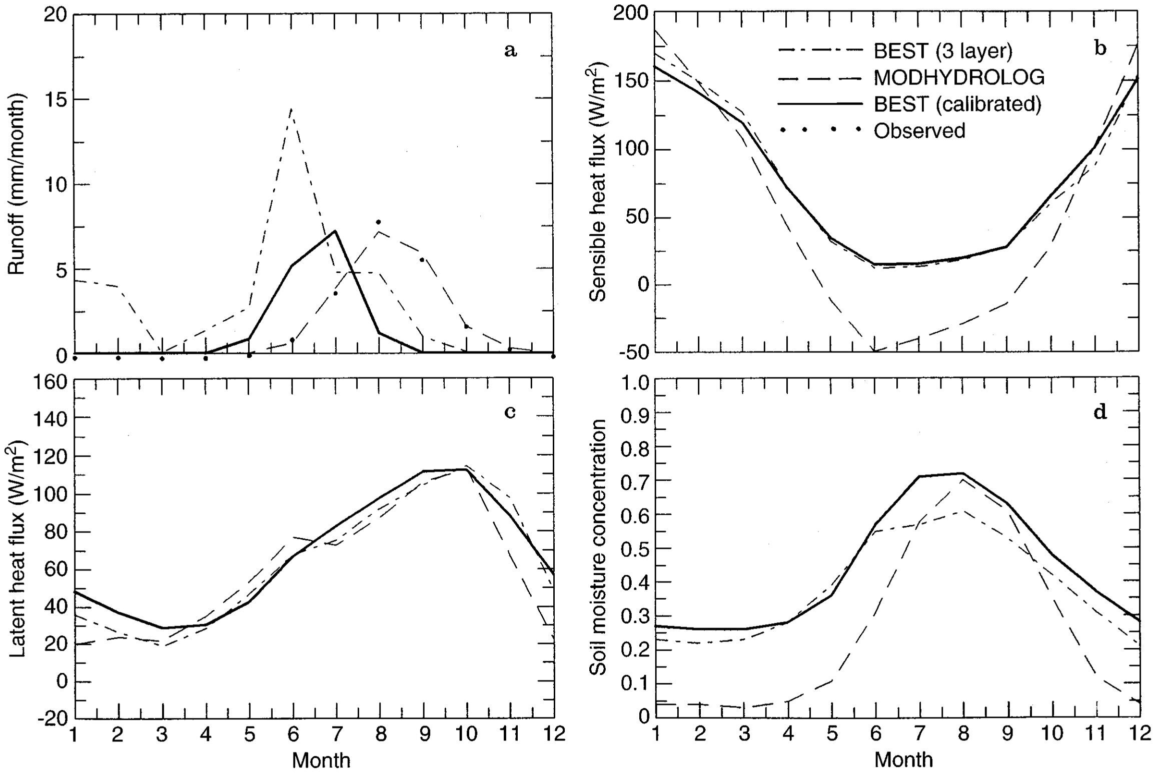

Fig. 11a–d. As Fig. 6 but for the three-layer version of BEST. The parameter values are $d_u = 0.1$ m and $d_l = 1.9$ m in the control run and $d_u = 1.9$ m and $d_l = 3.2$ m in the final simulation

7 Conclusions

A two-layer version of BEST was able to either simulate the observed runoff or simulate similar turbulent energy fluxes and soil moisture compared to MODHYDROLOG for a wet catchment (Jardine River). By increasing the soil moisture stores, the observed runoff was simulated realistically, but the simulation of the latent heat flux diverged from that simulated by MODHYDROLOG because the layer was too deep to dry and limit the latent heat flux. BEST proved unable to simulate the observed runoff for a drier, Mediterranean-type Canning River catchment.

The two-layer model was extended to a three-layer scheme which was able to simulate the observed runoff, and the latent heat flux and soil moisture simulated by MODHYDROLOG for the wet catchment. However, a good simulation of the Canning River catchment remained elusive.

It is suggested that three soil layers are a necessity for simulating both runoff and the latent heat flux realistically, at least for the Jardine River catchment. It is also suggested that heterogeneity in the runoff generation parametrisation is probably needed in order to simulate the Canning River catchment. The indication that different processes appear necessary in explaining the behaviour of the model in different catchments is not surprising but does indicate that developing a generic land surface scheme and applying it to all catchments (or grid squares) may prove difficult.

Finally, BEST has been tested against a variety of data sets of atmospheric flux data reasonably successfully prior to this work. This is the first time this model has been tested against runoff data. This proved to be a far more difficult test of the model. We concur with Arnell (1995) and Rowntree and Lean (1994) in stating that runoff data is extremely valuable in testing the performance of land surface models.

Acknowledgements. We thank the Australian Bureau of Meteorology, the Western Australia Water Authority and the Queensland Water Resources Commission for providing the climate and streamflow data. This project was carried out with the help of funds from the Australian Greenhouse Research Grants Scheme, the Australian Research Council and the Macquarie University Research Grants scheme.

Appendix A Brief description of MODHYDROLOG

Rainfall-runoff models are developed to estimate runoff from rainfall and potential evapotranspiration data. Focusing mainly on the movement of water, they pay little attention to the estimation of surface energy fluxes, except in the simulation of evapotranspiration.

Rainfall-runoff models are frequently classified into 'black box' or process models. In the 'black box' modelling approach, empirical equations are used to relate runoff and rainfall, and only the input (rainfall) and output (runoff) have physical meanings. Process models [e.g. the Systeme Hydrologique European (SHE) model, Abbott et al. 1986]; and Institute of Hydrology

(UK) Distributed Model (IHDM), Beven et al. (1987) simulate the hydrological processes in a catchment using partial differential equations governing various physical processes and equations of continuity for surface and soil water flow.

Most hydrological applications adopt a simpler approach. A catchment is conceptualised as a number of interconnected stores, with mathematical functions used to describe the movement of water between stores. These 'conceptual models' attempt to represent the physical processes but often include 'black box' treatment where empirical equations and 'effective' parameters are used to describe the processes.

The conceptual daily rainfall-runoff model used in this study, MODHYDROLOG, has been used to estimate streamflow particularly in Australia (see Chiew and McMahon 1994). The model structure, and the equations representing the various hydrological processes, are shown in Fig. A1 (model parameters are highlighted in bold). A detailed description of the model can be found in Chiew (1990), while Porter (1972) provides a detailed description of the origins of the equations used to represent the catchment hydrological processes.

In MODHYDROLOG, incident daily rainfall first fills the interception store, which is emptied each day by evaporation. The excess rainfall is then subjected to a function which determines infiltration. Some of the water which cannot infiltrate is diverted to a depression store (regulated by the depression flow function) while the remainder becomes surface runoff. The depression store empties by both evapotranspiration and delayed infiltration into the soil moisture store. All moisture that infiltrates is next subjected to a soil moisture function. This function diverts moisture to the stream as interflow and to the groundwater store as groundwater recharge. Moisture that is not diverted enters the soil moisture store. Evapotranspiration from the soil moisture store occurs at a rate which is dependent on the soil moisture status and potential evapotranspiration. The soil moisture store has a finite capacity and overflows into the groundwater store. The groundwater store can be depleted by baseflow into the stream and by deep seepage to the underlying aquifers or replenished by recharge from the stream and upwards movement of water from the underlying aquifers.

MODHYDROLOG takes into account spatial variation by allowing the user to apply the model individually to sub-areas within the one catchment (with different input data and parameter values). The outflow from each sub-area becomes inflow to the next sub-area, and together with the total runoff, is progressively routed to the catchment outlet using a non-linear routing technique. However, spatial variation is not allowed for in this study, and the total runoff (streamflow) simulated for the catchment is simply the sum of surface runoff, interflow and baseflow (see Fig. A1), routed from the catchment centroid to the outlet.

Potential evapotranspiration is specified and represents the atmospherically controlled upper limit of

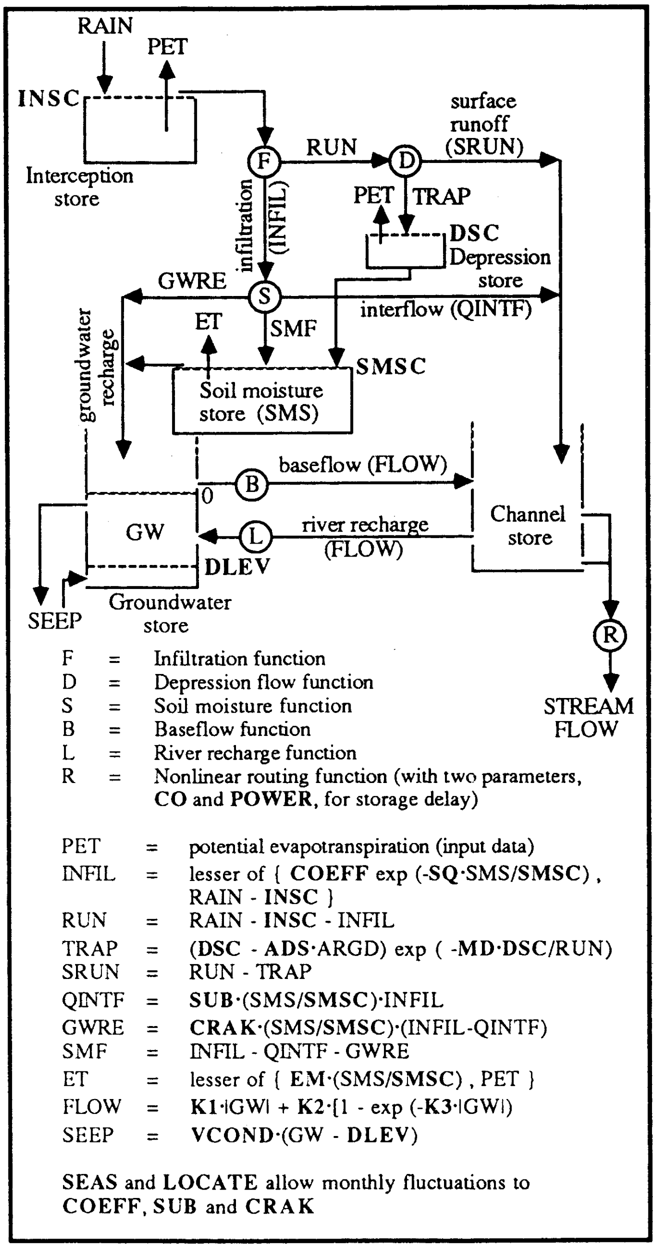

Fig. A1. Schematic representation of MODHYDROLOG

evapotranspiration from the soil moisture store. For this study, potential evapotranspiration is calculated following Morton (1983). The total evapotranspiration from MODHYDROLOG is the sum of evapotranspiration from the soil moisture store and intercepted rainfall (see Fig. A1). As the model uses a daily time step, it does not simulate the diurnal variation of evapotranspiration. Over a long time period the sensible heat flux is inferred as the difference between net radiation and the simulated latent heat flux.

The parameter values in MODHYDROLOG depend partly on the climate conditions and the catchment flow and physical characteristics (Porter and McMahon 1976; Chiew and McMahon 1994). As with other rainfall-runoff models, the parameters can vary widely and the correlations of some parameters with catchment characteristics can be poor. The models are therefore commonly calibrated by optimising parameter values to provide a good fit between the simulated and recorded flows. Optimisation of the 19 parameters to less than ten parameters is usually sufficient to give adequate estimates of streamflow, and the use of four or five parameters may be sufficient in temperate and wet catchments (Chiew and McMahon 1994).

References

Abbott MB, Bathurst JC, Cunge JA, O'Connell PE, Rasmussen J (1986) An introduction to the European Hydrological System – Systeme Hydrologique European, "SHE", 1. History and philosophy of a physically-based, distributed modelling system. J Hydrol 87:45–59

Abramopoulos F, Rosenzweig C, Choudhury B (1988) Improved ground hydrology calculations for global climate models (GCMs): soil water movement and evapotranspiration. J Clim 1:921–941

Arnell NW (1995) River runoff data for atmospheric model validation. In: Oliver HR, Oliver SA (eds) The role of water and the hydrological cycle in global change, NATO ASI Series 1 (Global Environmental Change Volume 31), Springer, Berlin Heidelberg New York pp 349–371

Beven KJ, Calver A, Morris EM (1987) The Institute of Hydrology Distributed Model, Institute of Hydrology, Wallingford, UK Rep 81

Bonan GB (1994) Land-atmosphere CO_2 exchange simulated by a land surface process model coupled to an atmospheric general circulation model. J Geophys Res 100:2817–2831

Budyko MI (1974) Climate and life. English Edn DH Miller (ed), Academic Press, New York, 508 pp

Chiew FHS (1990) Estimating groundwater recharge using an integrated surface and groundwater model. Ph. D thesis, Department of Civil and Agricultural Engineering, University of Melbourne, Victoria, Australia

Chiew FHS, McMahon TA (1994) The optimisation of parameters of the daily rainfall-runoff model MODHYDROLOG in 28 Australian catchments. J Hydrol 153:383–416

Chiew FHS, Pitman AJ, McMahon TA (1996) Conceptual catchment scale rainfall-runoff models and land-surface parameterisation schemes. J Hydrol, 179, 137–157

Cogley JG, Pitman AJ, Henderson-Sellers A (1990) A model of land surface climatology for general circulation models. Trent Tech. Note, 90-1, Trent University, Canada

Cosby BJ, Hornberger GM, Clapp RB, Ginn TR (1984)A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resour Res 20:682–690

Deardorff JW (1977) A parametrization of ground-surface moisture content for use in atmosphere prediction models. J Appl Meteorol 16:1182–1185

Deardorff JW (1978) Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation. J Geophys Res 83:1889–1903

Dickinson RE, Henderson-Sellers A, Kennedy PJ, Wilson MF (1986) Biosphere Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model. NCAR Techn Note, NCAR, TN275+STR, 69 pp

Dickinson RE, Henderson-Sellers A, Kennedy PJ (1993) Biosphere – Atmosphere Transfer Scheme (BATS) Version 1e as coupled to the NCAR Community Climate Model. NCAR Tech Note, NCAR, TN383+STR, 72 pp

Dolman AJ, Gregory D (1992) The parametrization of rainfall interception in GCMs, Q J R Meteorol Soc 118:455–467

Dümenil L, Todini E (1992) A rainfall-runoff scheme for use in the Hamburg climate model. In: O'Kane JP (ed), Advances in theoretical hydrology, A tribute to James Dooge, European Geophysical Society Series or Hydrological Sciences: 1. Elsevier

Gates WL, Rowntree PR, Zeng Q-C (1990) Validation of climate models. In: Houghton JT, Jenkins GJ, Ephraums JJ (eds). Climate change, the IPCC Scientific Assessment. Cambridge University Press, Cambridge, pp 93–130

Henderson-Sellers A, Dickinson RE, Durbidge TB, Kennedy PJ, McGuffie K, Pitman AJ (1993) Tropical deforestation.: modelling local- to regional-scale climate change. J Geophys Res 98:7289–7315

Jacobs CMJ, de Bruin HAR (1992) The sensitivity of regional transpiration to land surface characteristics: significance of feedbacks. J Clim 5:683–698

Mahrt L, Pan H (1984) A two-layer model of soil hydrology. Bound Layer Meteorol 29:1–20

Manabe S (1969) Climate and the ocean circulation: 1, The atmospheric circulation and the hydrology of the Earth's surface. Mon Weather Rev 97:739–805

Mihailovic DT, De Bruin HAR, Jeftic M, van Dijken A (1992) A study of the sensitivity of land surface parametrizations to the inclusion of different fractional covers and soil textures. J Appl Meteorol 31:1477–1487

Mintz Y (1984) The sensitivity of numerically simulated climates to land-surface boundary conditions. The global climate. Houghton JT (ed) Cambridge University Press, Cambridge, UK 79–105

Morton FI (1983) Operational estimates of actual evapotranspiration and their significance to the science and practice of hydrology. J Hydrol 66:1–76

Nobre CA, Sellers PJ, Shukla J (1991) Amazonian deforestation and regional climate change. J Clim 4:957–988

Noilhan J, Planton S (1989) A simple parametrization of land surface processes for meteorological models. Mon Weather Rev 117:536–549

Pitman AJ, Yang Z-L, Cogley JG, Henderson-Sellers A (1991) Description of bare essentials of surface transfer for the Bureau of Meteorology Research Centre AGCM. BMRC Research Report 32, 117 pp

Pitman AJ, Yang Z-L, Henderson-Sellers A (1993a) Sub-grid scale precipitation in AGCMS: re-assessing the land surface sensitivity using a single column model. Clim Dyn 9:33–41

Pitman AJ, Henderson-Sellers A, Abramopoulos F, Avissar R, Bonan G, Boone A, Cogley JG, Dickinson RE, Ek M, Entekhabi D, Famiglietti J, Garratt JR, Frech M, Hahmann A, Koster R, Kowalczyk E, Laval K, Lean L, Lee TJ, Lettenmaier D, Liang X, Mahfouf J-F, Mahrt L, Milly C, Mitchell K, de Noblet N, Noilhan J, Pan H, Pielke R, Robock A, Rosenzweig C, Running SW, Schlosser A, Scott R, Suarez M, Thompson S, Verseghy D, Wetzel P, Wood E, Xue Y, Yang Z-L, Zhang L (1993b) Results from the off-line control simulation phase of the Project for Intercomparison of Land surface Parametrisation Schemes (PILPS). GEWEX Techn note, IGPO Publication Series 7. 47 pp

Polcher J, Laval K, Dümenil L, Lean J, Rowntree PR (1995) Comparing three land surface schemes used in GCMs. Clim Dyn VV:PP–PP

Porter JW (1972) The synthesis of continuous streamflow. Ph. D Thesis, Department of Civil Engineering, Monash University, Victoria, Australia

Porter JW, McMahon TA (1976) The Monash model: user manual for daily program HYDROLOG. Department of Civil and Engineering, Monash University, Victoria, Australia, Res Rep 2/76, 41 pp

Rowntree PR, Lean J (1994) Validation of hydrological schemes for climate models against catchment data. J Hydrol 155:301–323

Sellers PJ, Dorman JL (1987) Testing the simple biosphere model (SiB) using point micrometeorological and biophysical data. J Clim Appl Meteorol 26:622–651
Sellers PJ, Mintz Y, Sud YC, Dalcher A (1986) A simple biosphere model (SiB) for use within general circulation models. J Atmos Sci 43:505–531
Sellers PJ, Shuttleworth WJ, Dorman JL, Dalcher A, Roberts JM (1989) Calibrating the simple biosphere model for Amazonian tropical forest using field and remote sensing data. Part I, average calibration with field data. J Appl Meteorol 28:727–759
Shao Y, Anne RD, Henderson-Sellers A, Irannejad P, Thornton P, Liang X, Chen TH, Ciret C, Desborough CE, Barachova O, Haxeltine A, Ducharne A (1994) Soil moisture simulation, a report of the RICE and PILPS Workshop. GEWEX Tech Note, IGPO Publ Series 14. 179 pp
Verseghy DL, McFarlane NA, Lazare M (1993) CLASS – a canadian land surface scheme for GCMs. II: vegetation model and coupled runs. Int J Climatol 13:347–370
Verseghy DL (1991) CLASS: a Canadian land surface scheme for GCMs, I. Soil model. Int J Climatol 11:111–133
Viterbo P, Illari L (1994) The impact of changes in the runoff formulation of a general circulation model on the surface and near surface parameters. J Hydrol 155:325–336
Warrilow DA, Sangster AB, Slingo A (1986) Modelling of land surface processes and their influence on European climate, Dynamic Climatology Tech Note 38, Meteorological Office, MET O 20. (unpublished), Bracknell, Berks., UK, 94 pp
Webb RS, Rosenzweig CE, Levine ER (1993) Specifying land surface characteristics in general circulation models: soil profile data set and derived water-holding capacities. Global Biogeochem Cycles 7:97–108
Whetton PH, Fowler AM, Haylock MR, Pittock B (1993) Implications of climate change due to the enhanced greenhouse effect on floods and droughts in Australia. Climate Change 25:289–317
Xue Y, Shukla J (1993) The influence of land surface properties on Sahel climate. Part 1: desertification. J Clim 6:2232–2245
Yang Z-L, Pitman AJ, McAvaney B, Henderson-Sellers A (1995) The impact of implementing the bare essentials of surface transfer (BEST) land surface scheme into the BMRC GCM. Clim Dyn 11:279–297