ELSEVIER

Global and Planetary Change 13 (1996) 57–71

GLOBAL AND PLANETARY CHANGE

Modeling vadose zone liquid water fluxes: Infiltration, runoff, drainage, interflow

Peter J. Wetzel [a,*], Xu Liang [b], Parviz Irannejad [c], Aaron Boone [a,d], Joel Noilhan [e], Yaping Shao [c], Chris Skelly [f], Yongkang Xue [g], Zong Liang Yang [h]

[a] *Mesoscale Atmospheric Processes Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA*
[b] *Department of Civil Engineering and Operations Research, Princeton University, Princeton, NJ, USA*
[c] *Climate Impacts Centre, Macquarie University, Sydney, NSW, Australia*
[d] *Science Systems and Appl., Inc., Lanham, MD, USA*
[e] *Meteo France / CNRM, Toulouse, France*
[f] *Department of Geography, James Cook University, Townsville, QLD, Australia*
[g] *Center for Ocean–Land–Atmosphere Studies, Calverton, MD, USA*
[h] *Institute of Atmospheric Physics, University of Arizona, Tucson, AZ, USA*

Received 5 May 1995; accepted 28 August 1995

Abstract

Because of the large water storage capacity of soil relative to the atmosphere, changes in soil moisture storage can significantly affect the regional atmospheric budgets of water and energy on monthly, seasonal and longer time scales. Therefore proper modeling of soil liquid water processes is essential to a correct representation of the climate system.

This study focuses on the class of summary models of liquid water fluxes in the vadose, or unsaturated zone of the soil, which are applicable to global or regional climate modeling studies. Fourteen such models are represented in this intercomparison study. Observational data from the HAPEX experiment provide validation. Because only limited observational data were available to constrain these models during their development and validation, the models have evolved very diverse treatments of the relevant processes: the basic Darcian (soil internal) and Hortonian (surface liquid flow) processes, as well as the boundary conditions of baseflow drainage and lateral interflow.

The annual total local runoff is systematically underestimated by all but one of the participant models. This is one of the few significant biases between the consensus of participant models and the observations. The modeled runoff, averaged over the 14 models, differs from the budget estimate from observations by about 40%. During the period of runoff generation (late winter and early spring) the average model fails to deplete the soil water store as rapidly as is observed, a result consistent with the underprediction of runoff. One cannot rule out insufficient characterization of the field site soils as a primary cause of these discrepancies. Results suggest that model sources of the discrepancy are about equally likely to be related to the prediction of bare soil evaporation (discussed elsewhere in this issue) as they are to the parameterization of runoff and drainage processes.

1. Introduction

This paper discusses simulation of the boundary conditions imposed upon the soil water domain within

* Corresponding author.

SSDI 0921-8181(95)00037-2

land surface parameterizations designed for implementation in three-dimensional atmospheric models. A physically based framework (the soil column model) is chosen for the basic structure of all of the participant models discussed in this study. However, most models have adopted conceptual models of the hydrologically-oriented sub-surface processes and boundary conditions in order to allow the column model to be applied to large-scale averages.

Two of the boundary conditions for which physical models are presumed to apply, evaporation from bare soil and plant extraction of water, are discussed in companion papers elsewhere in this issue. These two form the basis for direct interaction between soil liquid water processes and the atmospheric energy and water balances, and are therefore of special importance to atmospheric modeling. Since evaporation and local soil moisture are physically linked, the choice of a column model as the basic structure for parameterization in an atmospheric model is a natural one.

Nevertheless the sub-surface boundary conditions, those which pertain to liquid water flow, and ultimately determine stream flow, are of no less consequence for the successful long-term simulation of the atmosphere. Conceptualizing these sub-surface liquid water processes in a manner which efficiently transforms a "raw" single-column model into an effective and accurate "ensemble" soil column model is a serious challenge. Fourteen different solutions to this challenge are intercompared herein.

The soil liquid water store is generally at least two orders of magnitude larger than that held in the atmosphere. Soil spatial scales are compressed relative to the atmosphere, since there are no turbulent or advective processes involved, but time scales are large. Significant changes to the grid-averaged soil water storage occur on monthly to seasonal time scales. If just one of the subsurface fluxes which determine the rate of change of that store is inadequately posited, error in the predicted soil water available for evaporation will accumulate, over a period of weeks to months, to the point where serious systematic errors in the diagnosed fluxes result.

Given that all the sources and sinks of soil water in a basic hydrologic unit (a soil column, drainage basin or atmospheric model grid cell) are subject to measurement and modeling error, and that errors may randomly offset one another, the problem of establishing an optimal parameterization does not have a unique solution. Observational constraints on the fluxes of water vapor into the atmosphere are sufficient to allow some certainty in evaluation of evapotranspiration parameterizations. Similar constraints on the parameterization of the four major liquid water fluxes (overland flow, infiltration, interflow [lateral boundary flux], and drainage out the modeled soil lower boundary) do not exist. Unresolvable soil and topographic heterogeneity occurs within any practical atmospheric model's minimum resolution unit. The only readily available constraints on modeled liquid water processes at depth in the soil are the evapotranspiration measurements themselves, point and lysimeter measurements of soil moisture, measured point water table depth, perhaps estimates (via remote sensing) of the extent of saturated surface area and the extent of surface area below the wilting point, and the measured stream flow. The latter provides the only liquid flux measurement, and it provides only an integration (over multiple, indeterminate time and space scales) of all four liquid water flux processes combined. The available routes which a water molecule may take from the time it encounters the soil surface until it emerges at a stream gauging station, involve a virtually unknown network of surface channels and subsurface macropores and micropores which vary in size and conductive capacity by tens of orders of magnitude. It is no wonder, therefore, that the parameterizations of these four processes differ drastically from model to model.

Since none of the models participating in this study include an explicit representation of the water table, discussion will be limited to the vadose, or unsaturated soil zone. In the following section the model formulations are reviewed and limitations on the constraining observations are discussed. Then the results of intercomparison workshop experiments are presented and evaluated. The final section discusses conclusions and recommendations.

2. Model formulation

As implied above, soil column water modeling, the approach chosen by all participants, employs relatively simple theory, the implementation of which is almost entirely governed by empirical treatment of unspecifiable boundary conditions. In employing very simple boundary conditions, some of the models represented here have nearly entirely abandoned the theory. The most extreme example is the BUCKET model in which a soil layer simply absorbs all incident water until it becomes saturated. The next level of complexity is represented by the force-restore method, in which two layers of soil interact with one another through an empirically tuned pair of flux terms, one causing water to move in response to source and sink terms, and the other causing movement which adjusts toward a reference or base state. The ISBA model is an example of this. A third class of models attempts a low-resolution solution to the homogeneous medium diffusion equation. Among these is PLACE.

Fig. 1 provides a schematic summary of key hydraulic fluxes as modeled by the fourteen workshop participants. In terms of the number of modeled soil moisture layers, the models can be grouped into four classes. BUCKET and SECHIBA2 are one layer models with only one drainage process-a pseudosurface runoff; BEST, BIOME2, CSIRO, ISBA, and VIC are two layer models with flow out the sides and/or bottom; BGC, BATS, CLASS, LAPS and SSiB are three layer models with disparate boundary conditions; and CENTURY and PLACE have more than three layers, also with disparate boundary conditions. Most of the models explicitly include some sort of computation of the diffusion of liquid water between soil layers. Besides the one layer models, other exceptions to this are BGC, BIOME2 and VIC.

In terms of the drainage and subsurface flow formulations, the fourteen models can be classified into two major types: those keyed to values of intermediate soil water content (field capacity and/or wilting point) and those which apply a (continuous) nonlinear function relating soil moisture to key diffusive parameters such as hydraulic conductivity and water potential. A few models (such as BEST) actually fall into both categories. BEST computes diffusion, but keys its drainage to field capacity. On the

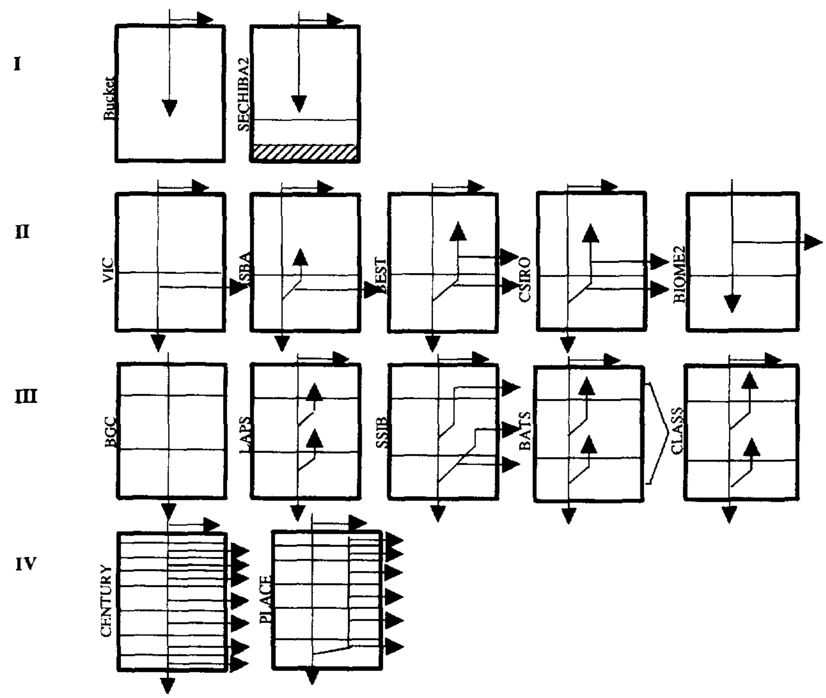

Fig. 1. Schematic illustration of runoff and drainage for land surface schemes represented in this study. According to the number of soil layers, the schemes are classified into group *I* (single layer models), group *II* (two layer models), group *III* (three layer models) and group *IV* (multi-layer models).

other hand, PLACE and VIC allow the soil to drain without limit, and asymptotically approach zero soil moisture during a very long dry period. The non-linear soil water functions (required to close the solution to the diffusion equation) vary widely from model to model, including empirical relationships derived from catchment hydrology, soil column studies, laboratory work and theory (see below for further discussion).

It should be noted that some models, originating in the hydrology discipline, use nonlinear, one dimensional empirical functions which are designed for an empirical fit to two and three dimensional subsurface flow fields (i.e. an entire catchment). These could benefit from a much more explicit hydrologic characterization of the HAPEX field site.

Further details and descriptions of the modeling of individual processes are presented in the following subsections.

2.1. Empirical representation of soil characteristic curves

The relationship between volumetric soil water content and soil hydraulic conductivity, diffusivity and water potential depend on the morphology of soil pores. Average pore size varies widely depending on soil type. Variability of pore sizes, if not stochastic, is usually neglected, despite the fact that heterogeneity can have a large and almost always systematic effect on soil water flux (Philip, 1980).

Simple, closed-form, homogeneous, one-dimensional solutions to the diffusion equation require continuous analytical functions to represent the relationship between soil water and conductivity or potential. The most popular relationship among the models represented here is the Brooks-Corey model (Brooks and Corey, 1966) which is a simple power law relationship (see Boone and Wetzel, this issue, for a more detailed discussion). VIC uses this model in its original form, which includes a parameter called "residual saturation", an amount of water that can never be extracted. Most other models apply the Clapp-Hornberger (1978) modification which is identical to Brooks-Corey except that it assumes zero residual saturation. Both of these models break down in the vicinity of saturation, rendering them perilous to use when representing the key process of infiltration (see Clapp and Hornberger, 1978). The well-known Green-Ampt formulation (Green and Ampt, 1911) offers a robust, independent treatment of infiltration. However, for numerical implementation, it carries the burden of requiring knowledge of the precipitation/infiltration history. On the other hand, the "bucket" assumption obviates the need for an infiltration equation. All incident water is assumed to be absorbed by a soil layer until it saturates.

An alternative to the Brooks-Corey model has been suggested by Van Genuchten (1980). In this model, Θ, defined as the available volumetric soil water content (the absolute volumetric water content minus the residual saturation value) nondimensionalized by (i.e. divided by) its saturation value, is related to similarly nondimensionalized soil water potential, Ψ by

$$\Theta = \left(\frac{1}{1 - \Psi^n} \right)^{1-2/n}$$

where n is an empirical parameter.

Recent analysis by Fuentes et al. (1992) compared various combinations of the Van Genuchten and Brooks-Corey models along with several others. They concluded that for practical application (i.e. in a land surface parameterization), the best of the eight models/combinations they tested was the combination of the Van Genuchten equation for Θ and the Brooks-Corey power law for normalized hydraulic conductivity, K:

$$K = \Theta^{2b+3}$$

where b is also an empirical parameter equal to $n-2$. This combination has not been tested in any of the models discussed here, and empirical data sets based on observed soils do not appear to have been published for this combination. It is mentioned here for informational purposes and to encourage further research. Fuentes et al. (1992) find this to be the only method they tested which "satisfies the infiltration condition for all soil types, even when applied to the two extreme cases used by Green and Ampt". This is a rigorous, empirical approach which applies simple, continuous, closed functions, and avoids any coding complications because it is universally applicable without restriction, to the entire range of soil moisture for all soil types.

2.2. Partitioning throughfall into infiltration and runoff

Throughfall, i.e. the portion of precipitation which is not intercepted or evaporated before it reaches the soil, is all absorbed into the soil until either of two conditions is met: the immediate surface soil becomes saturated and/or the local rate of throughfall exceeds the local infiltration capacity. The infiltration capacity is the maximum instantaneous rate of absorption of liquid water by the soil. In some models the surface soil may achieve saturation before the infiltration capacity drops to zero (i.e. there are deeper unsaturated layers which continue to absorb water). All water which is not instantaneously infiltrated ponds at the surface and begins to flow overland in response to topographic gradients. The surface water which is ponded or flowing overland may infiltrate later, or at a different, down-flow location, or it may enter permanent stream channels. For the purposes of this paper, runoff is defined as all water which is not instantaneously absorbed by the soil. None of the participant models explicitly track liquid surface water beyond the point where it is partitioned between infiltration and the various surface containment and flow processes; therefore, the latter are all lumped.

Further, it should be noted that none of the participant models explicitly or implicitly route runoff through a stream channel or subsurface reservoir, although other models in the wider community have included this feature. The function of a routing algorithm is to smooth and delay the instantaneous discharge. The result is a predicted runoff time series which more closely resembles measured stream gauge data.

Treatment of the partitioning between runoff and infiltration varies widely among the participant models. BATS routes a fraction of all precipitation a priori into runoff. BGC, CENTURY and BIOME2 do not permit explicit surface runoff. BUCKET, ISBA and most other force-restore models generate no runoff until soil saturation is achieved. BEST is notable because it applies a very simple form of the "contributing area" concept. It assigns a fraction of the surface to be saturated based on the spatial mean value of its top layer soil moisture. The saturated portion of the soil contributes to runoff while the remainder permits infiltration at a rate based on a solution of the soil water diffusion equation. The VIC model, by its very name (Variable Infiltration Capacity) also calculates an effective contributing area. The majority of the remaining models allow runoff only when it exceeds a single, spatially averaged infiltration capacity, usually calculated based on the diffusion equation. In most of these models, saturation excess runoff is an "all or nothing" process, not unlike a bucket model. The grid box is assumed to be either all saturated, or else none of it is represented as saturated.

2.3. Internal vertical flow between modeled layers

The flow of water within the soil is most commonly modeled using a form of the diffusion equation given by Richards (1931). This equation is divided into a purely diffusive term and a gravitational drainage term. In its practical application in one dimension, other source and sink terms, such as lateral discharge and extraction by evaporative processes, may be lumped with gravitational drainage to produce the equivalent of a "force" term, as in the simple force-restore model. The diffusive term is identical in concept to the "restore" term, although in the case of Richards equation, the equilibrium or reference value to which a layer is restored is not fixed, but is based on the weighted mean water content of adjacent layers. Thus the force-restore model can be seen as an empirically optimized special case of the one dimensional Richards equation. Because of the inherent heterogeneous, three dimensional nature of natural soils, the theoretical underpinning afforded Richards equation is not necessarily an advantage.

In any case, given the discussion above, one may lump the majority of participant models into one class—those applying some form or Richards equation. The exceptions are models which do not allow any diffusive processes, such as capillary rise, between layers. Among these are BUCKET, BGC, BIOME2, CENTURY, SECHIBA2 and VIC.

One caveat must be noted. The diffusion equation assumes that the medium is a homogeneous continuum. As a result, in virtually every form of finite numerical application, its solution unavoidably depends explicitly on the choice of model resolution. In the case of the relevant participant models here, the resolution is represented by the specified soil layer depth. Changing the model depth while holding everything else the same, will change the computed result, often significantly. (See Boone and Wetzel, this issue, for some further discussion.)

2.4. Drainage: the lower boundary condition

Once the lower boundary of a model is specified, the task remains to characterize the condition of the soil at that depth at the site to be simulated. A minimal characterization requires knowledge of the pore size and its distribution at the boundary. Grain size is not sufficient, since it provides no information about macropores. An adequate characterization also requires the time history of the soil liquid water potential at the boundary throughout the period to be modeled. If the model's lower boundary is near (within one meter of), or below the depth of the water table, this depth serves as a satisfactory proxy for the potential. Obviously, adequate characterization of even a small site is difficult without massive destructive sampling. When the model depth is large, or when the area to be represented is more than a few tens of square meters, true characterization becomes absolutely impossible. The experiments reported here rely on a very crude characterization. See the section on limitations of the observations (below) for further discussion of this.

All participant models except BUCKET, SECHIBA2 and BIOME2 allow some form of gravitational drainage out of the bottom of the model domain. In some cases, such as VIC and SSIB, the drainage out the bottom and the interflow (see below) are lumped together. Nearly all the drainage parameterizations have as their kernel the saturation hydraulic conductivity (which is equal to the rate of drainage that would occur under the force of gravity in the absence of any other potential gradient). Most models simply multiply the conductivity by a constant (less than 1) or by a simple function. In SSiB and some other models, this function is related to the topographic gradient. Boone and Wetzel (this issue) found that the PLACE model, which multiplies the conductivity by a fixed constant, is extremely sensitive to the value of the constant.

In an effective "ensemble" column model, which includes a sufficiently deep soil column (well below the root zone), the obvious way to parameterize drainage is by using baseflow (i.e. rain-free period) recession curves from stream gauge observations. This is a naturally integrated measure which is relatively easy to extract from streamflow records. The parameters of such a recession model could be considered site-specific in the same way that vegetation type is. There may be just a few classes of parameter choices which can be applied globally.

2.5. Interflow: lateral discharge from a soil layer

Where the soil surface is sloped, a significant possibility exists that differential horizontal subsurface flow of liquid water may occur. About half of the participant models take this into account. The hydraulic conductivity of the layer generally serves as the kernel. A few of the models (e.g. PLACE) identify one or two parameters which are explicitly tied to the topography of the domain. The remaining models with lateral discharge make no pretense as to the physical significance of the required empirical parameters. There is some debate in the community as to whether physically tying sub-surface flow parameters to surface topography makes any sense. Perhaps on a point-by-point basis it makes less sense than it does in a statistically aggregated model.

2.6. Limitations of the observations which constrain the models

Although the goal of this workshop was to focus on the models, some insight into the observations is helpful in interpreting the results presented here and elsewhere in the issue. Despite the fact that the carefully collected HAPEX–MOBILHY data is truly an outstanding accomplishment, and is among some of the best data available, it necessarily suffers some limitations.

The Caumont site (see Goutorbe et al., 1989), used for the forcing data and for validation in this study, was instrumented with a single flux measurement station with a homogeneous fetch of a few hundred meters. Recent results from the FIFE [First ISLSCP (International Satellite Land Surface Climatology Project) Field Experiment] site, reported by Grossman (1995), indicate that the "footprint" of a single surface station's latent heat flux measurement may be significantly larger than 200 m. Using a sophisticated, physically based (rather than purely statistical) separation technique, Grossman (1995) found that a substantial fraction of the measured near-surface water vapor flux is contributed by wavelengths of from 200 m to at least 800 m, and is contained in eddies which have their origin at the boundary layer top. Therefore, he concludes that consideration of a much wider area is required in order to correctly characterize the footprint of the measured surface flux. Surface heterogeneities further upwind of the Caumont flux station may explain the apparent inconsistency between the evapotranspiration estimated from analysis of the site's liquid water budget and that measured by the flux station (see below and other papers in this issue).

Measured soil characterization was limited to a sampling of grain size, presumably of the near-surface soil. This, along with the neutron probe soil moisture data, which were averaged from four individual measurements made within 10 m of the flux station, comprise the only local soil information available. Runoff and water table depth were not measured, and no information on soil heterogeneity, either in the horizontal or the vertical was provided. Examination of the individual neutron probe measurements reveal an average standard deviation among the four simultaneous measurements, of usually between 1 and 2% by volume below 20 cm depth. This provides some indication of the heterogeneity of the soils on the 10 m spatial scale, and is very typical of values reported in the literature for many other sites. Greater standard deviations in the near surface measurements reflect significant measurement error inherent in the neutron probe method.

The site was agricultural—a soya bean field. The site was bare during the winter, covered by nothing more than crop stubble, and was tilled at about day 120. The effect of tilling is to significantly change the soil pore morphology of the top 15 or 20 cm of soil, and it can actually decrease the permeability below, by closing macropores. No attempt was made in this study to represent a sudden change in soil characteristics associated with tilling.

A number of the flux measurement sites within HAPEX-MOBILHY were irrigated, including the one used in this study (Caumont). Goutorbe et al. (1989) report that all the irrigated sites except Caumont were irrigated with sprinklers. Irrigation water was captured in the station rain gauge at these sites, although significant error can be related to the non-vertical trajectory of droplets and an uneven distribution of water by the sprinkler system. At Caumont, unfortunately, the irrigation amount was not measured, but simply estimated. During the period from 28 May to 18 July, 30 mm of rain fell at Caumont and 24 mm of irrigation is assumed—an amount which perfectly closes the water budget. In the forcing data used in this study, irrigation water is imposed on two occasions (14 mm at sunset on 3 July and 10 mm at noon on a nearly cloud free day—18 July), Since these amounts are assumed rather than measured, doubt remains as to the error limits to assign the measured water budget.

2.7. Some concluding remarks on model diversity

Given the discussion in the preceding subsections, a picture emerges of a largely unconstrained subsurface water budget. Significantly more well-constrained data do not exist except for highly instrumented small hydrologic research watersheds (perhaps one of these should be identified for future study). The models participating in this study were developed within this unconstrained environment, largely to serve the purposes of the atmospheric or ecological modeler (with but one exception). Therefore the use of many diverse means to reach the same end (namely a good representation of the local or regional, daily or seasonal flux of water vapor) should not surprise the reader.

3. Results and discussion of the experiments

The goal of the analysis presented below is to identify the degree of difference among models in the total predicted liquid water loss, and to attempt to explain these differences. Analyses of two experiments are presented: The control experiment, in which each model is allowed to freely seek its own equilibrium, and a special "pulse precipitation" experiment in which evaporation is turned off in order to isolate model differences in soil water physics. In the control experiment, most attention will be devoted to the first four months of the calendar year, before the beginning of the growing season. During this period soil water was observed to oscillate about the apparent field capacity. Soil liquid water flow processes are perceived to dominate over the small amount of bare soil evaporation which took place during this period. On the other hand, during the peak of the growing season, plant extraction of water dominates, and drainage becomes a secondary process.

3.1. Control experiment—the partitioning of precipitation into liquid water loss and evaporation

Results from the most recent control runs, experiments 13 and 15 are shown in Figs. 2 and 3, respectively. These experiments are discussed in more detail elsewhere in this issue. Briefly, Experiment 13 represents the latest (presumably error-free) control run following the strict guidelines for surface characterization specified by PILPS. In experiment 15, modelers were allowed to optimize any specialized model-specific parameters which were not specified by PILPS. For example, the difference between

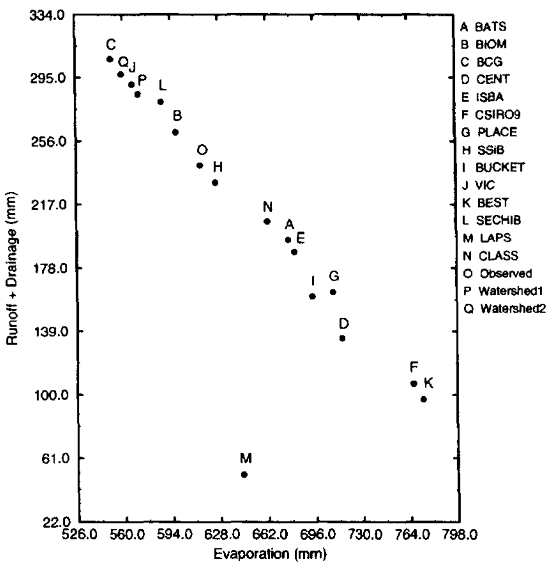

Fig. 2. Comparison of annual total runoff plus drainage against evaporation (mm) for different schemes from Experiment 13. For the data point represented by the symbol *O*, the evaporation is estimated from the Penman-Monteith equation.

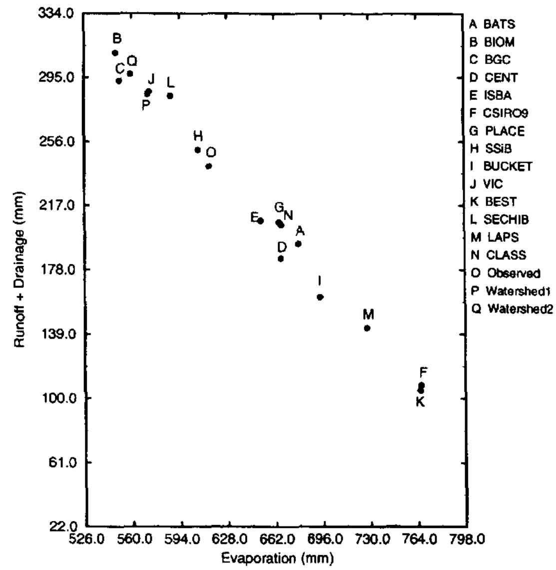

Fig. 3. Same as Fig. 2 but for Experiment 15.

experiment 13 and 15 for the PLACE model is that no soil heterogeneity is allowed in Experiment 13, but in experiment 15, three soil types, averaging to the specified one, were represented (see further discussion in section 3.3 below. A number of models did not submit results for Experiment 15. Their results are reported as identical for the two experiments.

In Figs. 2 and 3, the annual total liquid water loss (hereafter called runoff and drainage) is plotted against the annual total model predicted evaporation. Since these are equilibrium results, no change in soil water storage should occur, and all models should fall on a single line representing the observed total precipitation. Clearly the differences between models is large. Annual total drainage ranges from about 100 to 300 mm. Unfortunately, no runoff data were collected during the field study. A comparison of observed and simulated runoff time series could shed light on important model characteristics. There were, however, two small catchments on either side of the study site, from which ten years of runoff data have been analyzed by Goutorbe et al. (1989). Based on the ten year average of the ratio of runoff to precipitation, and assuming, without justification, that subsurface water flow characteristics of the study site are similar, each catchment can provide a rough "observed" value. In Figs. 2 and 3, these are indicated as watershed1 (285 mm) and watershed2 (297 mm). Another "quasi-observed" point is plotted as well. It is derived from a budget which applies a simple Penman-Monteith calculation of evaporation, with parameters empirically optimized at the study site in order to fit the available observations.

Although points are rather uniformly strewn along the diagonal total precipitation axis in Figs. 2 and 3, several groupings of models can be defined: First, a minority of the models (BIOME2, BGC, SECHIBA, SSiB and VIC) fall within the range of the observations. The other nine models appear to underestimate total drainage. Of these, six cluster around a total drainage of about 185 mm. Without regard to cause and effect, one can generalize that models which under predict drainage are left with too much available water for evaporation. BGC and BIOME2 use the field capacity threshold approach in their subsurface flow and drainage formulations; VIC uses an empirical nonlinear function of soil water for subsurface flow and drainage based on large scale catchment hydrology; SSiB generates runoff from its saturated soil layer and from its nonlinear base flow lower boundary condition. It is difficult to conclude, based on the annual runoff results, which model structure or theory is preferable.

Based on the soil moisture at the beginning of the equilibrium year, the models can be divided into two groups: (1) those starting at soil moisture close to the observed (BATS, BGC, BIOME2, CSIRO9, ISBA and VIC are within 10 mm) and (2) the remainder of the models, which begin the year with soil moisture lower than observed. It seems that the models which simulate the soil moisture content well in the dry period, June to October, tend to also recover to the correct observed soil moisture by the beginning of January. The models which take too much moisture via transpiration during the dry period naturally take longer to recover in the fall and early winter.

Examination of the 14 models' annual time series of drainage reveals that the vast majority of each models' annual total drainage occurs between days 0 and 120, i.e., before the onset of the growing season. Therefore explanations for the scatter in Figs. 2 and 3 can be sought by examining each model's soil water budget during this period.

3.2. Estimated soil water budget—January through April

An approximate water budget for days 0–120 can be generated using the observed weekly root zone (0–1.6 m) soil moisture, accumulated precipitation, and the empirically site-adjusted Penman-Monteith estimate of evaporation. Evaporation is calculated at 149.6 mm for the period, amounting to about 70% of net radiation. The total precipitation was 368.5. The observations show very little change in total soil moisture during the period. Water content of the soil layer from 0.5 to 1.6 m shows virtually no change, while the upper half meter holds steady during January and February then drops during March and April, presumably in response to the increased demand for bare soil evaporation. The total root zone water content change is estimated to be −22.2 mm between days 0 and 120. One further inference can be made about the water budget: Since the maximum precipitation rate (12 mm/h) did not exceed the estimated saturation hydraulic conductivity (14.4 mm/h) we can infer that (1) the soil surface never became saturated, and (2) little infiltration excess runoff could have occurred. Therefore all runoff is likely to have passed through subsoil channels. The numbers given above yield a total of 241.1 mm of runoff generated during the first 120 days of the year. Note that nothing can be said about the rate of realization of this runoff as streamflow.

In stark contrast with the estimate based on observations, all fourteen models' control runs predicted an increase in soil moisture between day 0 and day 120. BGC, BIOME2, CENTURY, ISBA and PLACE are the models which have the smallest increase in soil moisture (less than 30 mm) during this time. BATS, CENTURY and CLASS have soil moisture increases of less than 50 mm, and the rest of the models have soil moisture increases of more than 50 mm with SSiB the highest (more than 110 mm). Based on the drainage and subsurface flow formulation classified earlier, it is seen that BGC, BIOME2, CENTURY, and ISBA are the models which have the field capacity threshold formulation. However, only BGC and BIOME2 drain more than 200 mm (close to the amount obtained from the observed soil water budget), ISBA and CENTURY drain less than 160 mm. They both are compensated by high evaporation to achieve the small soil moisture change during this period. PLACE simulated a closer soil moisture change for the first 120 days compared with the observation. Although PLACE has a deep drainage formulation which takes care of the presence of the macropore or fracture flows, it seems that it still does not drain fast enough to obtain the small change in soil moisture without the compensation from its bare soil evaporation (more than 200 mm). BATS gives a small change in soil moisture during this period due to its large bare soil evaporation (more than 220 mm) and its surface runoff (89.6 mm). The total runoff plus drainage from BATS is less than 120 mm for the first 120 days. In fact, all of the models showed that the surface runoff (both excess infiltration runoff and saturation runoff) was very small (less than 5.0 mm) which seems to be realistic. The one exception, BATS, generates such a large surface runoff due to assigning a proportion of precipitation as a priori surface runoff.

Consistent with the overprediction of soil moisture change, all schemes underestimated drainage during the first 120 days. The one exception to this was BGC, which simulated slightly higher runoff (but within the 10% error range). Fig. 4 shows the time series of predicted runoff by each model during the period. BIOME2, ISBA, SECHIBA2 and VIC simulated the runoff larger than 150 mm. The simulated runoff from the rest of the schemes is about half of the observed 241.1 mm, with BEST and CSIRO9 the lowest (less than 80 mm). BGC, BIOME2, and ISBA have steeper slope in the first 60 days of accumulated runoff and drainage than the rest of the models, and most the models have steep slopes in the accumulated runoff from day 95 to day 120, except for BEST, CSIRO9, and LAPS.

A plot of each model's drainage error (difference from the observed 241.1 mm) during the 120 days vs. its change in predicted soil moisture reveals an interesting grouping of models. Fig. 5 shows that nine of the models fall within a relatively narrow envelope along a line passing through the observed point at lower right, with a slope of −0.55. This implies that these nine models agree that about 55% of the difference in their predicted soil moisture is due to the inter-model differences in predicted drainage, and the remaining 45% of the disparity in predictions of soil moisture is due to model differences in predicted bare soil evaporation. These percentages could be case specific. However, at least in this one case, they imply some consistency among model formulations concerning the relative impor-

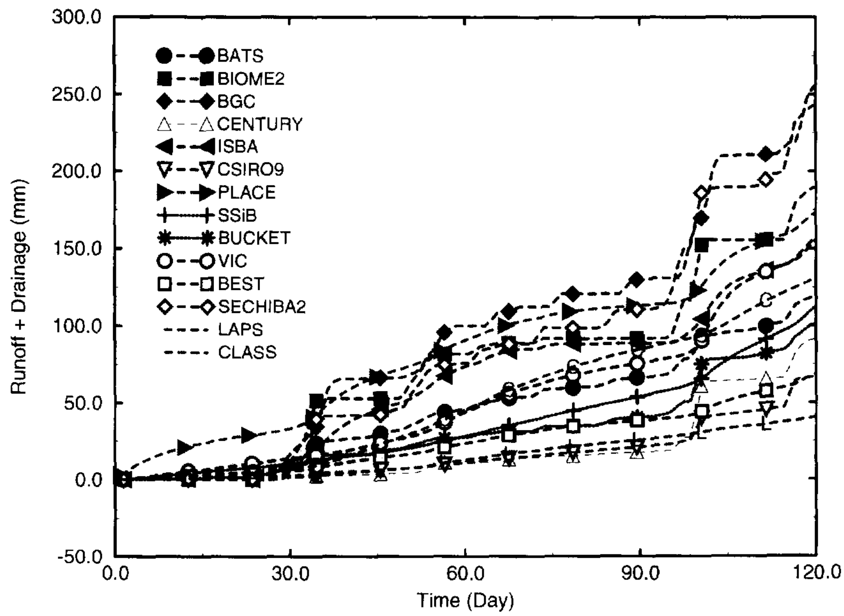

Fig. 4. Integrated runoff plus drainage (mm) for the first 120 days from Experiment 13 (or Experiment 15, when available).

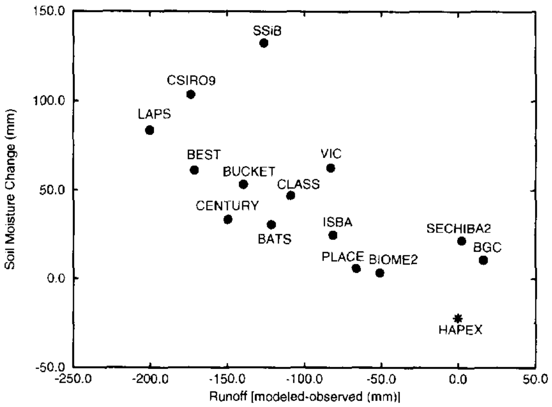

Fig. 5. Comparison of soil moisture change (mm) against the difference between predicted and observed runoff for Experiment 13 (or Experiment 15, when available).

tance of evaporation and drainage sub-models to the soil moisture prediction. The two are roughly equally important. Therefore both must be given careful consideration. If the models, which all lie skewed to one side of the observed point, are to converge toward the observations, carefully balanced, simultaneous modifications (because of the physical feedbacks) to both the drainage and evaporation submodels will be required.

3.3. Hypothesized cause of the systematic drainage error: soil heterogeneity

Since all models but BGC require greater drainage in order to come into closer agreement with observation, an examination of neglected processes that would increase drainage is in order. Philip (1980), in working a one dimensional problem with soil heterogeneities in series, showed that drainage in such a system is restricted by the least conductive inclusion (layer). He also discussed the fact that heterogeneity in parallel is essentially additive—that is, each adjacent soil column conducts water as if it were homogeneous. What his consideration of parallel heterogeneity failed to account for, however, is the statistical bias toward lateral (horizontal) transfer of water from less porous to more porous zones, both at the surface and within the vadose zone. At the surface, the likelihood exists that surface ponding and overland flow will occur more quickly and more frequently over less conductive soils than over adjacent more porous soils. Some of the overland flow may encounter porous zones where it can be infiltrated. Below the surface, a more porous soil will fill and drain more quickly as a result of a finite rainfall event than its adjacent less porous soil. After an initial relatively brief period in which the more porous soil is wetter at any level, a more lengthy period will follow in which the less porous soil remains wetter. Thus the horizontal moisture gradients which occur will statistically favor horizontal flow toward the more porous area over the long term. The result is that the resultant permeability (and drainage) of a soil system with horizontal heterogeneity, i.e., heterogeneity in series, will approach the permeability of the most porous constituent soils. Put simply, in the natural soil system, whether or not surface processes are included, horizontal heterogeneity systematically increases drainage relative to homogeneous soils with identical mean properties.

A plethora of observations exist demonstrating a high degree of observed horizontal variability in soils even within a single field. Given these points, it would seem that consideration of soil heterogeneity, including representing the effect of macropores may be a promising direction for future work. As one example, the PLACE model experiment 15 result included three explicit soil types, equally weighted, whose average clay, silt and sand contents equal the observed value (silt content held constant, clay and sand percentages varied by $+11/0/-11$) Further, within each of the three soil types, a normal distribution of soil water content was assumed, with variability defined by a 2%-by-volume standard deviation. The agreement between this model result and observations improved noticeably compared to PLACE experiment 13 result with just a single soil type. Comparing point G on Figs. 2 and 3 reveals that PLACE's annual evaporation was significantly improved. The error in annual runoff production was reduced by 50%. At the same time, PLACE's prediction of evaporation during the Intensive Observation Period (considered within the envelope of observational error) changed very little between runs 13 and 15. Perhaps most importantly, PLACE run 15 predicts root zone (0–1.6 m) soil moisture very close to the observed, whereas run 13 was significantly too dry. Annual mean root zone soil moisture, observed to be 434.2 mm, is predicted to be 421.4 mm using heterogeneous soils (run 15), but 382.5 mm when a homogeneous soil is assumed (run 13). (Full graphical results for the soil moisture annual cycle, not shown here for the sake of brevity, may be found elsewhere in this issue, and are available by ftp from pilps@cic.mq.edu.au.)

3.4. Interaction of the bare soil evaporation submodels with the model predicted absolute value of soil moisture

Fig. 6 shows that the difference between modeled and observed evaporation increases with a decrease in the soil moisture content at the beginning of the equilibrium year, with exceptions of SECHIBA2 and SSiB. This could be interpreted to mean that models which predict higher rates of bare soil evaporation generally achieve a lower winter equilibrium value of soil moisture. A line with a slope of −0.5 could comfortably be drawn among the points, as in Fig. 5, however the scatter here is much greater. Thus there are much greater model-to-model differences in the predicted equilibrium absolute value of soil moisture than there are in the predicted time change thereof.

Six models in Fig. 6 appear to start the calendar year with about the observed soil moisture, yet only BGC and SECHIBA end up with about the same soil moisture on day 120 with about the same total runoff and evaporation; BIOME2 and PLACE achieve nearly the same soil moisture at an expense of a

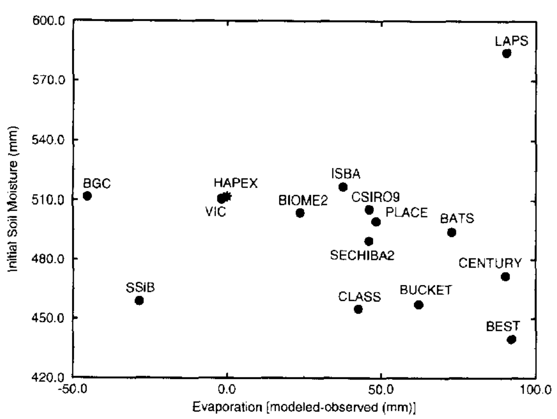

Fig. 6. Comparison of initial soil moisture (soil moisture at the beginning of the year) vs. the difference between predicted and observed evaporation (mm).

relatively large bare soil evaporation. Although VIC simulated bare soil evaporation close to observed, it reaches higher soil moisture than observed because it has less total runoff during this period. The smaller total runoff during the first 120 days from VIC is due to little hydrologic characterization information for the HAPEX field site in determining the time delay parameters in its empirical representation of subsurface flow and drainage. CSIRO9 produced high bare soil evaporation (more than 200 mm), it still ended up with higher soil moisture due to its very small total runoff (less than 70 mm). The rest of the models that finish with about the correct soil moisture on day 120 are from the category that start with soil moisture lower than observed.

From the brief analysis discussed earlier, it is seen that the effects of the drainage formulations of each model are highly variable from one to another. Since the actual sub-soil boundary conditions are not available (indeed never available) from the observations, it is difficult to compare anything but the net result, in this case an inferred instantaneous observed runoff, with model predictions of the same. Further complicating matters is the fact that drainage effects always interact with the effects from the initial soil moisture and evaporation.

3.5. The pulse precipitation drainage experiment

In order to isolate quantitative model differences in predicted drainage, an experiment was devised which limits the effects of initial soil moisture content and bare soil evaporation. A "pulse precipitation" experiment was the result (experiment 2c1). In this experiment the precipitation is specified 10 times larger than the observed precipitation for the first 60 days, then set to be zero from day 61 to day 120. During the 120 days, the bare soil evaporation is set to be zero. The initial soil moisture content is set at $0.05\ m^3/m^3$.

Based on the available results from eight models, it is seen that seven models (BATS, BEST, BUCKET, CSIRO9, PLACE, SSiB, and VIC) have similar patterns and magnitudes of total runoff and drainage. ISBA simulates about 630 mm of total runoff, CSIRO9 has about 750 mm, BATS, BUCKET, SSiB and VIC simulate about 850 mm, while BEST and PLACE are at about 900 mm (Fig. 7). The difference

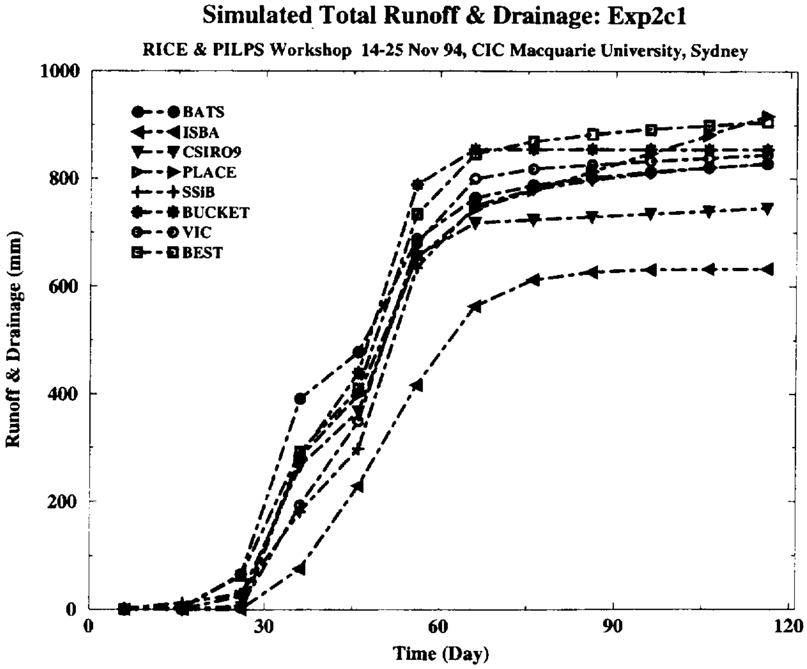

Fig. 7. Comparison of the integrated runoff plus drainage (mm) for the first 120 days from Experiment 2c1.

in total runoff and drainage among the seven models is about 150 mm in experiment 2c1, while it was about 100 mm in experiment 13 (or 15).

Fig. 8 shows the soil moisture variations among the eight models for the 120 days. The soil moisture of all the schemes approaches field capacity (512

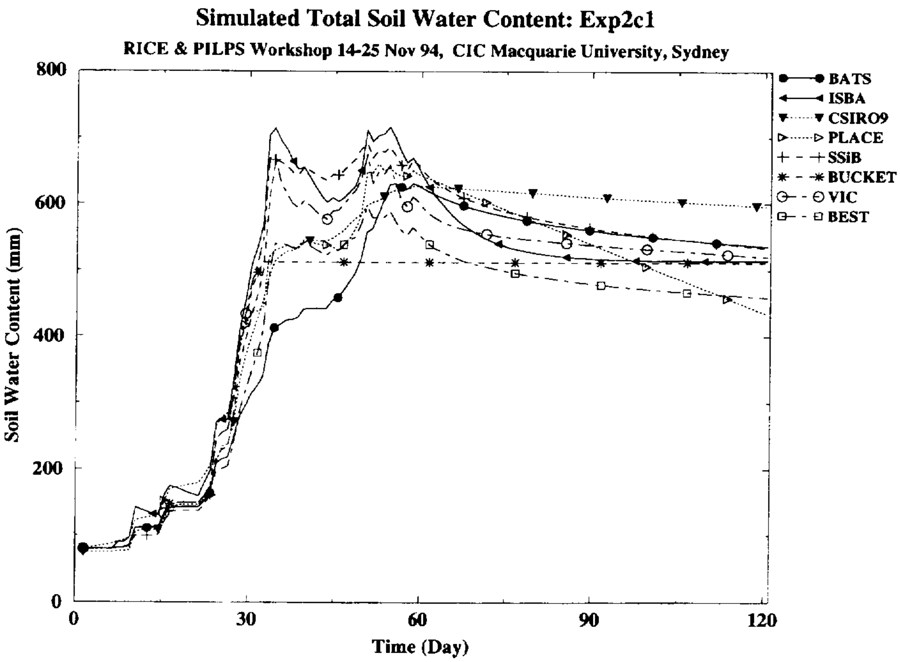

Fig. 8. Soil moisture variations from Experiment 2c1 among eight models for the first 120 days.

mm) on day 120, except for BEST, CSIRO9 and PLACE. Note that BEST uses a different effective field capacity than the one specified by the workshop, toward which it converges. PLACE continues to slowly drain until all water leaves the soil. In addition, the soil moisture reaches the field capacity around day 30 in all schemes, except for BATS which does not reach field capacity until about day 50. This is because of its larger surface runoff (Fig. 7).

All of the seven models exhibit quite a flat slope in the accumulated total runoff and drainage after day 65 except PLACE, which has steeper slope. In experiment 2c1, the percentage differences in runoff and drainage among the models are reduced compared with the results from experiment 13 (or experiment 15), based on only eight models' results. The results from experiment 2cl suggest that the large differences among model runoff from experiment 13 (or experiment 15) for the first 120 days may be mainly due to the differences in initial soil moisture content, bare soil evaporation, and the drainage and subsurface flow from unsaturated soil conditions, and less from the saturation subsurface flow and drainage.

4. Conclusions and recommendations

The combination of the model structure, the number of soil moisture layers and the runoff and drainage formulations lead to the large differences in the accumulated annual total runoff and in the accumulated runoff of the first 120 days. The simulation of total runoff and drainage appears to be about equally dependent on the simulation of soil moisture content and evaporation, as shown above. The water partition between total evaporation and total runoff and drainage shows that all but four of the schemes underestimate the annual total runoff and drainage and thus overestimate annual total evaporation. the comparisons also show that the majority of the schemes, which have too much water taken by transpiration from the soil system during the growing season, tend to start with smaller soil moisture contents than the observed in their equilibrium year. However, all the schemes simulate realistic surface runoff at the HAPEX site except for BATS which predicts very large surface runoff.

A "pulse precipitation" experiment was designed to study the difference in total runoff and drainage induced by the difference in model structures and formulations. There is some indication from the observed data that there may be some kind of rapid subsurface and drainage flow (such as selective flow through more porous zones in heterogeneous soils, macropore flow and pipe flow) at the HAPEX site. Therefore, any models that consider spatial soil variability in their runoff and drainage formulations should perform better here, although care must be taken to simulate the evaporation adequately. Also, experiment 2cl indicates that the percentage differences in the runoff and drainage are reduced among seven models reporting results for experiment 2cl compared with differences among the same seven models in experiments 13 (or 15). In addition, Experiment 2cl indicates that the schemes show better agreement in the saturation subsurface flow and drainage.

As there was no observed runoff time series at the HAPEX site, the performance of each model cannot be evaluated very satisfactorily. It would be useful to have observed runoff time series at the studied site in future evaluation and intercomparison projects.

Finally, the obvious issues of scale and relevance of the single surface site vis-a-vis a grid-averaged surface parameterization must be mentioned. The scale of the validation data in this study is clearly much smaller than the intended application of the models being tested. Perhaps even more importantly, too much emphasis on results from a single site can distract from the requirement for universality of the parameterizations, which are meant to apply globally and even under conditions which may not exist in the present day climate.

Acknowledgements

The dynamic and inspired leadership of Dr. Ann Henderson-Sellers is most gratefully acknowledged. Ann's ability to grasp the crux of scientific issues and separate the relevant detail from the trivial is uncanny. Despite the breakneck forward pace she maintains, there is rarely a misstep. Her ability to

simultaneously listen, focus and lead is indeed a rare gift. The entire staff of the Climate Impacts Centre deserves enormous credit for organizing and maintaining a most productive workshop—an enterprise which had its beginnings well before the actual convocation, and continues unabated to this writing.

References

Brooks, R.H. and Corey, A.T., 1966. Properties of porous media affecting fluid flow. J. Irrig. Drain. Am. Soc. Civ. Eng., IR 2: 61–88.

Clapp, R.B. and Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. Water Resour. Res., 14: 601–604.

Fuentes, C., Havercamp, R. and Parlange, J.-Y., 1992. Parameter constraints on closed-form soilwater relationships. J. Hydrol., 134: 117–142.

Goutorbe, J.-P., Noilhan, J. Valancogne, C. and Cuenca, R.H., 1989. Soil moisture variations during HAPEX-MOBILHY. Ann. Geophys., 7: 415–426.

Green, W.H. and Ampt, G.A., 1911. Studies on soil physics, Part I: The flow of air and water through soils. J. Agric. Sci., 4: 1–24.

Grossman, R.L., 1995. Length-scale composition of moisture flux in a continental, convective boundary layer. In: FIFE 2nd Spec. Issue. J. Geophys. Res., in press.

Philip, J.R., 1980. Field heterogeneity: Some basic issues. Water Resour. Res., 16: 443–448.

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. Phys., 1: 318–333.

Van Genuchten, M.T., 1980. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J., 44: 892–898.