Title

Key results and implications from phase 1(c) of the Project for Intercomparison of Land-surface Parametrization Schemes

Journal

Climate Dynamics, 15(9)

ISSN

0930-7575

Authors

Pitman, AJ
Henderson-Sellers, A
Desborough, CE
et al.

Publication Date

1999-09-01

DOI

10.1007/s003820050309

Peer reviewed

Climate Dynamics (1999) 15:673–684

A. J. Pitman · A. Henderson-Sellers
C. E. Desborough · Z.-L. Yang · F. Abramopoulos
A. Boone · R. E. Dickinson · N. Gedney · R. Koster
E. Kowalczyk · D. Lettenmaier · X. Liang
J.-F. Mahfouf · J. Noilhan · J. Polcher · W. Qu
A. Robock · C. Rosenzweig · C. A. Schlosser
A. B. Shmakin · J. Smith · M. Suarez · D. Verseghy
P. Wetzel · E. Wood · Y. Xue

Key results and implications from phase 1(c) of the Project for Intercomparison of Land-surface Parametrization Schemes

Received: 15 October 1997 / Accepted: 22 April 1999

Abstract Using atmospheric forcing data generated from a general circulation climate model, sixteen land surface schemes participating in the Project for the Intercomparison of Land-surface Parametrization Schemes (PILPS) were run off-line to equilibrium using forcing data from a GCM representative of a tropical forest and a mid-latitude grassland grid point. The values for each land surface parameter (roughness length, minimum stomatal resistance, soil depth etc.) were provided. Results were quality controlled and analyzed, focusing on the scatter simulated amongst the models. There were large differences in how the models' partitioned available energy between sensible and latent heat. Annually averaged, simulations for the tropical forest ranged by 79 $W m^{-2}$ for the sensible heat flux and 80 $W m^{-2}$ for the latent heat flux. For the grassland, simulations ranged by 34 $W m^{-2}$ for the sensible heat flux and 27 $W m^{-2}$ for the latent heat flux. Similarly large differences were found for simulated runoff and soil moisture and at the monthly time scale. The models' simulation of annually averaged effective radiative temperature varied with a range, between all the models, of 1.4 K for tropical forest and 2.2 K for the grassland. The simulation of latent and sensible heat fluxes by a standard 'bucket' models was anomalous although this could be corrected by an additional resistance term. These results imply that the current land surface models do not agree on the land surface climate

A. J. Pitman (✉) · C. E. Desborough
Dept. Physical Geography, Macquarie University,
North Ryde, 2109, NSW, Australia
E-mail: apitman@penman.es.mq.edu.au

A. Henderson–Sellers
Australian Nuclear Science and Technology Organisation,
Sydney, Australia

Z.-L. Yang · R. E. Dickinson
Institute of Atmospheric Physics, University of Arizona,
Tucson, USA

F. Abramopoulos · C. Rosenzweig
Science Systems and Applications Inc., New York, USA

A. Boone · J. Noilhan
Meteo–France, Toulouse, France

N. Gedney · J. Smith
Hadley Centre, Bracknell, England

R. Koster
Hydrological Sciences Branch, NASA/GSFC, Greenbelt, USA

E. Kowalczyk
CSIRO Division of Atmospheric Research, Aspendale, Australia

D. Lettenmaier
Dept. Civil Engineering, University of Washington, Seattle, USA

X. Liang · E. Wood
Princeton University, Princeton, USA

J.-F. Mahfouf
ECMWF, Reading, England

J. Polcher
Laboratoire de Meteorologie Dynamique du CNRS, Paris,
France

W. Qu
RMIT, Melbourne, Australia

A. Robock
Rutgers University of New Jersey, New Brunswick, USA

C. A. Schlosser
COLA, Calverton, USA

A. B. Shmakin
Institute of Geography, Moscow, USA

M. Suarez
NASA/GSFC, Greenbelt, USA

D. Verseghy
Climate Research Branch, Atmospheric Environment Service,
Downsview, Canada

P. Wetzel
Mesoscale Dynamics and Precipitation branch,
NASA/GSFC, Greenbelt, USA

Y. Xue
Department of Geography, University of Maryland,
College Park, USA

when the atmospheric forcing and surface parameters are prescribed. The nature of the experimental design, it being offline and with artificial forcing, generally precludes judgements concerning the relative quality of any specific model. Although these results were produced de-coupled from a host model, they do cast doubt on the reliability of land surface schemes. It is therefore a priority to resolve the disparity in the simulations, understand the reasons behind the scatter and to determine whether this lack of agreement in de-coupled tests is reproduced in coupled experiments.

1 Introduction to the Project for the Intercomparison of Land-surface Parametrization Schemes (PILPS)

1.1 Overview of PILPS

To begin the process of comparing land surface models used in general circulation models (GCMs) and numerical weather prediction (NWP) models, the WMO-CAS Working Group on Numerical Experimentation (WGNE) and the Science Panel of the GEWEX Continental-scale International Project (GCIP) launched the Project for the Intercomparison of Land-surface Parametrization Schemes (PILPS). A key objective of PILPS (Henderson-Sellers et al. 1995) is to achieve greater understanding of the capabilities of land-surface schemes in atmospheric models. In an attempt to realize this objective, Phase 1 of PILPS was initiated where a series of land surface models were forced 'offline' with data generated from a GCM and results from the final year of a multi-year equilibrium simulation were reported and analyzed.

Early results from phase 1(a) of PILPS (Pitman et al. 1993) demonstrated that a set of land surface models partitioned available energy between latent and sensible heat fluxes with a large degree of difference. The simulation for a tropical forest grid point ranged by 90 $W m^{-2}$ for the sensible heat flux and 80 $W m^{-2}$ for the latent heat flux. For a grassland grid point simulations ranged by 44 $W m^{-2}$ for the sensible heat flux and 37 $W m^{-2}$ for the latent heat flux. These results led to a revision of the PILPS experimental design and to the construction of a series of supplemental experiments and consistency tests which ensured all models: converged to a steady state; conserved energy and water on the annual mean; and used the correct precipitation forcing. Sixteen schemes passed these tests, and these models comprise phase 1(c) of PILPS. Details of these schemes can be found in Table 1. Note that there are two models derived from work by Budyko (1956) including the standard Manabe (1969) 'bucket' hydrology model ($BUCK^C$) and $GFDL^F$ which follows Manabe (1969) but includes a simple stomatal-type resistance formulation in the calculation of evaporation.

This study will briefly report on phase 1(c) of PILPS by first reviewing the experimental framework and then discussing equilibrium simulation results at annual and seasonal time scales for the 16 schemes. This provides background and context for the results described by Koster and Milly (1997) who report results from additional experiments conducted as part of phase 1(c).

1.2 Framework of PILPS experimental design

All simulations conducted in phases 1 and 2 of PILPS have been conducted 'offline' (i.e. no atmospheric feedbacks) since it is easier to isolate the reasons for different results in offline simulations. In phase 1 of PILPS, each participant was provided with one year of atmospheric data (at 30 min resolution) obtained from a GCM for grid points representative of tropical forest (centred on 3 °S, 60 °W hereafter FOREST), a Northern Hemisphere mid-latitude grassland grid point (centred on 42 °N, 255 °W hereafter GRASS), and a tundra grid point (not discussed here). Modelling groups could interpolate or aggregate these data to a time interval of their choice using a provided program. The atmospheric data comprised downward shortwave (solar) radiation, downward infrared radiation, precipitation, air temperature, wind speed and specific humidity. In the case of FOREST there was high downward solar radiation throughout the year (monthly averages of 200–250 $W m^{-2}$). There was considerable seasonality in the precipitation with rainfall exceeding 600 mm in February and October, and falling below 150 mm in May through to August. In the GRASS case there was significant seasonality on the solar forcing with a minimum of 75 $W m^{-2}$ in January and a maximum of about 250 $W m^{-2}$ in May and June. The rainfall was less than 100 mm $month^{-1}$ except in July and August when it increased to about 300 mm and 450 mm respectively. The monthly mean air temperature was below freezing in January, February and December (with precipitation falling as snow), and reached 300 K in July.

All the simulations described here used *identical* atmospheric forcing. Since these data were taken from a GCM they are not necessarily realistic for a particular location or the prescribed vegetation type and so it is inappropriate to attempt to validate these simulations against observational data. Phase 2 of PILPS attempts to validate a set of land surface schemes against observational data (e.g. Chen et al. 1997).

A list of surface parameters were supplied to each PILPS group who were asked to characterize their land surface as closely as possible to these parameters. The aim was to try to ensure that differences in the simulations resulted from the parametrizations included in each scheme rather than differences in the albedo, roughness length, soil porosity or other prescribed parameters. We attempted to ensure that Table 2 was

Table 1 A list of those models and primary contacts who have provided data to PILPS. Some basic information describing each scheme is included.

Key	Model	Contact	Number of canopy layers	Interception treated	Number of layers for				Methodology used		Reference
					T[a]	Θ[b]	Roots	Canopy	Temperature	Soil moisture	
A	BASE	C.E. Desborough A.J. Pitman	1	Yes	3	3	2	Aerodynamic	Heat diffusion	Philip-de Vries	Desborough and Pitman (1998)
B	BATS1 E	R.E. Dickinson Z.-L. Yang	1	Yes	2	3	2	Penman/ Monteith	Force-restore	Darcy's Law	Dickinson et al. (1986, 1993)
C	BUCK	A. Robock A. Schlosser	0	No	1	1	1	None	Heat balance	Bucket + variation	Robock et al. (1995)
D	CLASS	D. Verseghy	1	Yes	3	3	3	Penman/ Monteith	Heat diffusion	Darcy's Law	Verseghy (1991); Verseghy et al. (1993)
E	CSIRO9	E. Kowalczyk J.R. Garratt	1	Yes	3	2	1	Aerodynamic	Heat diffusion	Force-restore	Kowalczyk et al. (1991)
F	GFDL	P.C.D. Milly	0	No	1	1	1	None	Heat balance	Bucket + variation	Milly (1992)
G	GISS	F. Abramopoulos C. Rosenzweig	1	Yes	6	6	6	Aerodynamic	Aerodynamic	Darcy's Law	Abramopoulos et al. (1988)
H	ISBA	J. Noilhan J-F. Mahfouf	1	Yes	2–3	2	1	Aerodynamic	Force-restore	Force-restore	Noilhan and Planton (1989)
I	MOSAIC	R. Koster M. Suarez	1	Yes	2	3	2	Aerodynamic	Force-restore	Darcy's Law	Koster and Suarez (1992)
J	PLACE	P. Wetzel A. Boone	1	Yes	7	5	2	Aerodynamic and Ohm's law[1]	Heat diffusion	Darcy's Law	Wetzel and Chang (1988) Wetzel and Boone (1995)
K	SECHIBA2	J. Polcher K. Laval	1	Yes	7	2	2	Aerodynamic	Heat diffusion	Choisnel	Ducoudre et al. (1993)
L	SPONSOR	A.B. Shmakin	1	Yes	1	2	2	Aerodynamic	Heat balance	Bucket + variation	Shmakin et al. (1993)
M	SSIB	Y. Xue	1	Yes	2	3	1	Penman/ Monteith	Force-restore	diffusion	Xue et al. (1991)
N	UGAMP	N. Gedney	1	Yes	3	3	2	Aerodynamic	Heat diffusion	Darcy's Law	Gedney (1995)
O	UKMET	J. Smith	1	Yes	4	4	4	Penman/ Monteith	Heat diffusion	Darcy's Law	Warrilow et al. (1986) Gregory and Smith (1994)
P	VIC-3L	Xu Liang E. Wood D. Lettenmaier	1	Yes	2	3	3	Penman/ Monteith	heat-diffusion	Philip-de Vries[c]	Liang et al. (1994; 1996)

[a] Temperature, [b] soil moisture, [c] also treats horizontal water flow through the soil

The aerodynamic method is used for a canopy evaporating at the potential rate. A separate Ohm's law analog for water stressed conditions', dependent on plant and soil internal potentials and resistances was used

Details on the basic formulations used for the canopy can be found in Garratt (1992); those on the soil temperature calculations in Deardorff (1977) for those schemes which use versions of the force-restore model or Hillel (1982) for those models using the heat diffusion method. The soil moisture parametrizations show less consistency: several groups use the force-restore (Deardorff 1977) or the Philip-de Vries method (Philip and de Vries 1957; de Vries 1958).

complete but it did not prove possible to set all land surface parameters in all land surface schemes identically. For brevity, this list only includes those parameters which were used by a majority of the land surface schemes. Full details of the recommended surface characterization and the atmospheric forcing data can be obtained from the authors or from http://www.cic.mq.edu.au/pilps-rice. All models were initialized for January 1st, all moisture stores (including the canopy interception store) were initialized as 50% of full capacity whether liquid or frozen. Snow mass was initialized as zero and all soil or canopy temperatures were initialized at 300 K for FOREST and 275 K for GRASS. Each model was run to equilibrium which was defined as when a scheme's January surface temperature changed between year n and $n + 1$ by less than 0.03 K, and the root zone soil moisture changed by less than 1 mm.

In the following, we do not attempt to identify the 'right answer' since the atmospheric forcing and choice of parameters were not intended to reproduce any particular geographical location. Therefore, terminology used in the following analysis such as 'anomalous' does not imply an incorrect result, only a simulation which appears to fall outside a consensus.

Table 2 List of parameters used in the PILPS experiments

Parameters independent of vegetation type	
Height of the lowest model level (m)	45
Near Infrared albedo of fresh snow	0.65
Visible albedo of fresh snow	0.85
Thermal emissivity for all surfaces	1.0
Fraction of precipitation coverage	1.0
Topography height (m)	0.0
von Karman constant	0.378
Bare soil aerodynamic roughness length (m)	0.01
Snow aerodynamic roughness length (m)	0.00024
Bucket depth (m)	0.15
Critical soil moisture (m)	0.1125
Depth of top soil layer (m)	0.1
Total soil depth (m)	10.0
Root distribution (fraction of total roots in top soil layer)	0.80
Minimum soil suction (m)	0.2
Snow albedo	0.75

Parameters dependent on vegetation type	GRASS	FOREST
Surface albedo (snow free)	0.22	0.138
Budyko's runoff coefficient	0.2	0.6
Maximum fractional cover of vegetation	0.8	0.9
Vegetation aerodynamic roughness length (m)	0.1	2.0
Displacement height (m)	0.0	18.0
Minimum stomatal resistance ($s\,m^{-1}$)	200.0	150.0
Maximum leaf area index	2.0	6.0
Minimum leaf area index	0.5	5.0
Stem area index	4.0	2.0
Light dependence of stomatal resistance ($m^2\,W^{-1}$)	0.02	0.06
Rooting depth (m)	1.0	1.5
Soil porosity	0.51	0.6
Maximum hydraulic conductivity ($mm\,s^{-1}$)	0.45×10^{-2}	0.16×10^{-2}
Fraction of water content at which permanent wilting occurs	0.378	0.487
Clapp and Hornberger "B" parameter	6.8	9.2
Ratio of soil thermal conductivity to that of loam	0.95	0.80
Canopy visible albedo	0.1	0.04
Canopy near infrared albedo	0.3	0.20

2 Annually averaged results

2.1 Annual statistics for the tropical forest

Figure 1a shows how annually averaged net radiation is divided between sensible and latent heat for each land surface scheme for the tropical forest simulation. Figure 1a shows that all models except BUCK^C simulated an annually averaged latent heat flux between 108 $W\,m^{-2}$ and 135 $W\,m^{-2}$ (equivalent to 10.6% of the annual precipitation, 3267 mm, or 31% if BUCK^C is included). The relatively high latent heat fluxes simulated by UKMET^O and BUCK^C were due to the specification of surface (or stomatal) resistance. In the case of UKMET^O the surface resistance was a single value which was fixed at the prescribed value of the minimum stomatal resistance (Table 2) which was unlikely to be a reasonable value for the UKMET^O model given that many models simulate higher resistances during the day. In the case of BUCK^C the stomatal resistance to the latent heat flux was not included which led to an anomalous latent heat flux compared to the other models. The difference between BUCK^C and GFDL^F was therefore due to the inclusion of stomatal resistance in the latter model. All except two models (UKMET^O and BUCK^C) simulated annually average sensible heat flux between 7 $W\,m^{-2}$ and 25 $W\,m^{-2}$.

In Fig. 1a, scatter along a line drawn from the top left to the bottom right of the figure is caused by differences in the partitioning of net radiation between latent and sensible heat. This was caused by a combination of all the thermal and hydrological processes included in each land surface scheme. Scatter away from this line indicates differences in net radiation and it can be seen that there is greater agreement for net radiation than there is for its partitioning between sensible and latent heat. There are two reasons for net radiation differences. First, some scatter was due to albedo differences simulated by those schemes which incorporated the Sellers et al. (1986) parametrization of canopy albedo. In this parametrization, albedo is derived (and not prescribed) in a complex way which is difficult to set up

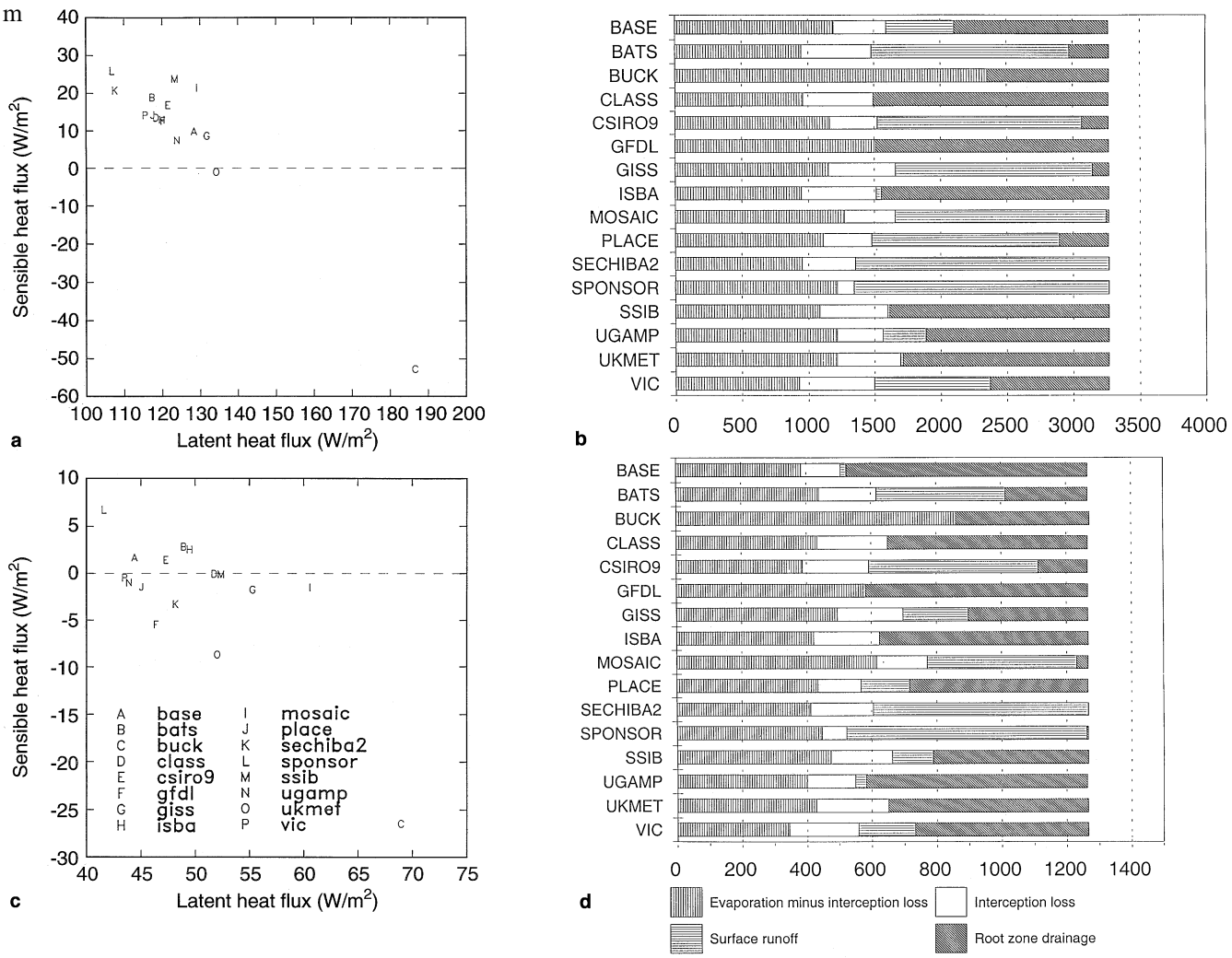

Fig. 1 a Relationship between annually averaged simulated sensible and latent heat fluxes from the models participating in PILPS for FOREST. **b** Water balance quantities for each model for FOREST (in mm y^{-1}). Note that root zone drainage for ISBA includes 1041 mm for root zone drainage and 675 mm for interflow. **c** As **a** but for GRASS. **d** As **b** but for GRASS

to produce an *a priori* albedo. Therefore, $SSIB^M$ and $MOSAIC^I$ simulated lower albedo and a higher net radiation than all the other schemes, and this was reflected in marginally higher latent and sensible heat fluxes. Second, scatter was caused by variations in upward infrared radiation resulting from differences in the calculated surface temperature (surface emissivity was set to unity).

The predicted annually averaged surface temperatures lay between 299 K ($MOSAIC^I$ and 300.4 ($SECHIBA2^K$). Five models ($MOSAIC^I$, $GISS^G$, $SSIB^M$, $CSIRO9^E$ and $BASE^A$ were below 299.5 K and four models were warmer than 300 K (VIC^P, $SPONSOR^L$, $SECHIBA2^K$ and $CLASS^D$). This range in temperatures led to differences in the upward infrared flux which explains much of the scatter perpendicular to the main axis of scatter in Fig. 1a. The two schemes with a high latent heat flux ($UKMET^O$ and $BUCK^C$) simulated an annually averaged effective surface temperature close to the median value of all the schemes (not shown).

The correct partitioning of precipitation between evaporation and runoff is an important reason for including sophisticated land surface schemes in GCMs and NWP models. Figure 1b shows how the land surface schemes partition the prescribed precipitation between evaporation (excluding evaporation of water intercepted by the canopy), root zone drainage, surface runoff and canopy evaporation (of intercepted water). Figure 1b shows that, for FOREST, the range in the simulated total annual evaporation was between 2353 mm ($BUCK^C$) and 1346 mm ($SPONSOR^L$). Some schemes which parametrize surface runoff did not simulate any ($CLASS^D$) while other schemes simulated no or negligible root zone drainage in the annual mean ($SECHIBA2^K$, $MOSAIC^I$, $SPONSOR^L$). Since the models were in equilibrium, the total range in predicted surface runoff plus root zone drainage was the same as the range in total evaporation, but the difference in the partitioning of runoff between root zone drainage and surface runoff was considerable. For

interception loss (evaporation of intercepted water) there is a range from 129 mm (SPONSOR^{L}) to 531 (BATS^{B}) which is 3.9% to 16.3% of precipit ation. This is a considerable range given that the interception capacity was prescribed (0.1 mm times the leaf area index). SPONSOR^{L} appears to be anomalous simulating only 129 mm of canopy evaporation in FOREST.

2.2 Annual statistics for the grassland

The grassland simulations were qualitatively similar to those for FOREST. Figure 1c shows the relationship between the simulated annually averaged sensible and latent heat fluxes for grassland (GRASS). The total scatter for GRASS is about 34 $W m^{-2}$ for sensible heat flux and about 27 $W m^{-2}$ for the latent heat flux. BUCK^{C} simulated a large negative (i.e. downwards) sensible heat flux but the addition of a stomatal resistance term to BUCK^{C} produced a reasonable simulation (e.g. GFDL^{F}). Overall, the range in the latent heat flux is 26% of the total precipitation (1268 mm). If BUCK^{C} and MOSAIC^{I} are omitted, the range is reduced to 12% which is similar to the range found for FOREST. The anomalous net radiation simulated by MOSAIC^{I} is partly related to the albedo parameter values provided by PILPS.

The range in the surface temperature simulated for GRASS was larger than for FOREST, with GFDL^{F} being coldest (280.1 K) and VIC being warmest (282.3 K). The larger total range (2.2 K) was expected given the larger seasonality in the atmospheric forcing and more complex nature of the surface (with snow, more bare soil (20% versus 10%)).

The partitioning of precipitation in the GRASS simulation is shown in Fig. 1d. The range in simulated total evaporation was from SPONSOR^{L} (522 mm) to BUCK^{C} (861 mm), a range which equates to 27% of the prescribed rainfall, but even if BUCK^{C} is omitted, the range remains 265 mm (20% of prescribed rainfall).

In terms of the interception loss (evaporation of intercepted water), SPONSOR^{L} was again anomalous simulating only 75 mm (Fig. 1d). The overall range was from SPONSOR^{L} to UKMET^{O} (223 mm) which, expressed as a percentage of precipitation, represented a range in interception loss from 5.9% to 17.6%.

2.3 Summary to annual statistics

At first sight, the scatter in the latent heat flux, and indeed temperature, appears quite large. However, if outliers are omitted, twelve of the sixteen models are within 15 $W m^{-2}$ for FOREST and thirteen of the sixteen models are within 10 $W m^{-2}$ for GRASS. Given the range in the complexity of parametrizations included in land surface schemes and uncertainty in the ability to characterize the surface parameters identically between schemes, an annual scatter of order 10 $W m^{-2}$ is probably as good as could be expected. The current performance of land surface schemes at the annual time scale therefore gives us some confidence that most of these models are capturing aspects of the partitioning of available energy between sensible and latent heat.

Overall, there was little general tendency for specific models to be warmer or cooler than other models (i.e. a model which simulated FOREST as relatively warm cold, rarely simulated GRASS also as warm cold). SECHIBA2^{K} and VIC^{P} are two exceptions, simulating both temperatures as quite warm while MOSAIC^{I} simulated both temperatures as quite cold. Similarly, a model which simulated an anomalous quantity for FOREST need not simulate the same quantity anomalously for GRASS. Again, some exceptions to this were SPONSOR^{L} which simulated relatively low evaporation and low interception and UKMET^{O} which simulated relatively high evaporation, but in neither case were the simulations very different from the broad consensus. The similarity in the overall interception loss between FOREST and GRASS was perhaps surprising given the different nature of the surface, the magnitude, intensity and duration of the precipitation and the differences in interception capacity.

These broad conclusions do not apply to BUCK^{C} which partitioned available energy between sensible and latent heat differently to any other model. BUCK^{C} also partitioned available water between runoff and evaporation in a very different fashion compared to the other schemes. The addition of a stomatal resistance term (e.g. GFDL^{F}) appears to remove these anomalies at the annual time scale. The anomalous behaviour by BUCK^{C} may be related to the use of an artificial drag coefficient which was different to the value of this coefficient typically used by the model. The role of the specification of this coefficient is discussed in more detail by Chen et al. (1997).

3 Seasonal results for tropical forest

3.1 Energy fluxes

Figure 2a shows the simulation of net radiation by each model. Although the parameters used by each model were prescribed, the different albedo parametrizations and problems with providing comparable parameter values for all types of model led to some seasonal variability in the albedo and absorbed solar radiation (not shown). In the case of monthly albedo, all models simulated albedo between 0.12 and 0.14 except MOSAIC^{I} and SSIB^{M} which both used a canopy albedo parametrization similar to Sellers et al. (1986). The parameter values provided for these schemes were

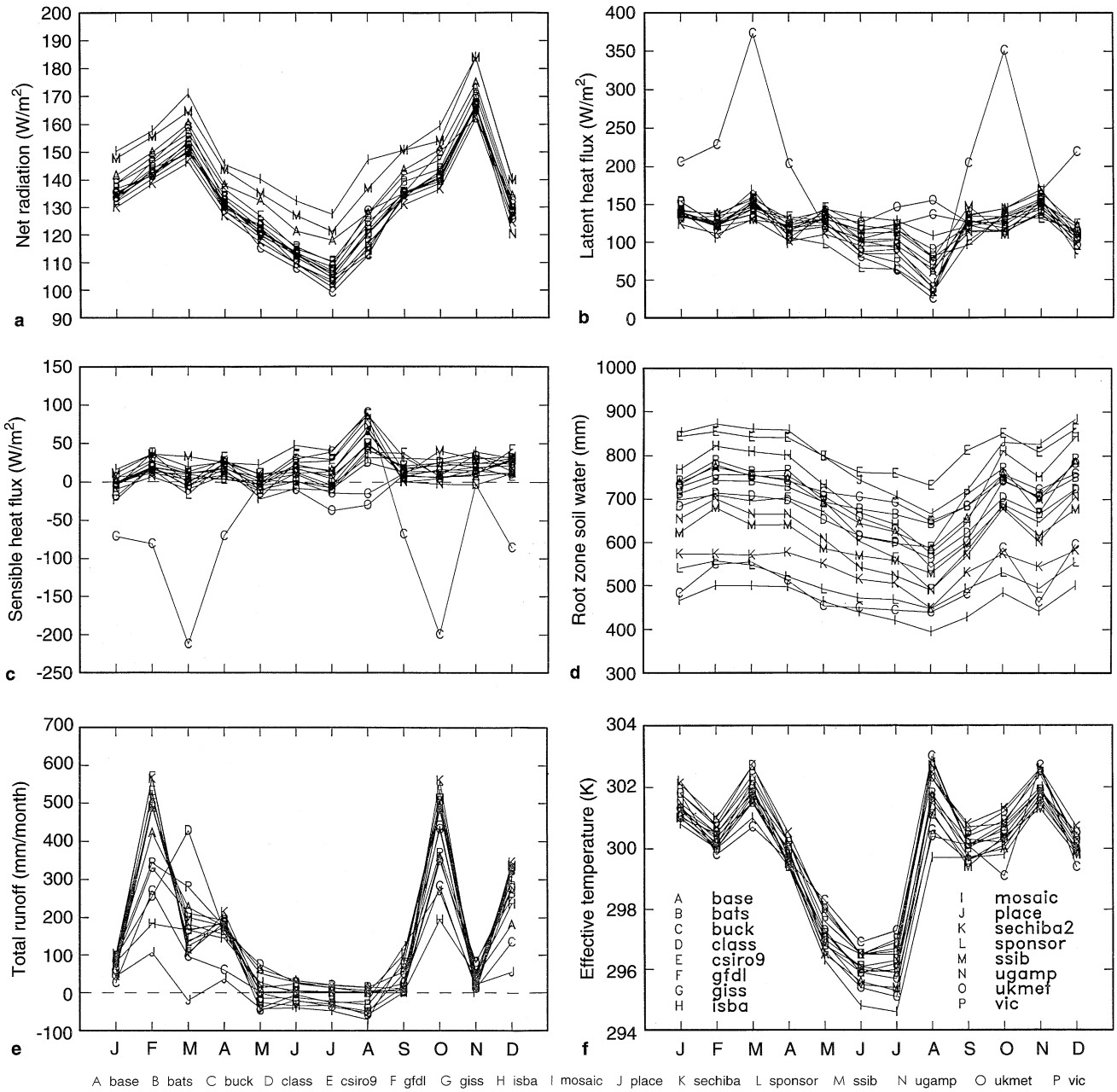

Fig. 2 a Mean monthly net radiation for FOREST; **b** as **a** but for the latent heat flux; **c** as **a** but for the sensible heat flux; **d** as **a** but for root zone soil moisture; **e** as **a** but for total runoff; **f** as **a** but for effective surface temperature

unrealistic for a tropical forest, hence the resulting albedos were too low leading to higher net radiation by 10–20 $W m^{-2}$. Three models which prescribe the albedo ($UGAMP^N$, $ISBA^H$ and $BUCK^C$) use an albedo of 0.138 (Table 2), but since most of the models predict albedo, these three schemes simulate 3 $W m^{-2}$ less absorbed solar radiation. While the use of different albedos must affect the simulation of other quantities, the effects were small and do not change any of the overall conclusions. Overall, the variability in the net radiation (Fig. 2a) was about 20 $W m^{-2}$ in January and 35 $W m^{-2}$ in July. The absorbed solar radiation (not shown) explained about 15 $W m^{-2}$ of this scatter, the remaining scatter was due to differences in the infrared exchange.

The differences in the net radiation (due to the differences in effective temperatures and albedos) led to variations in the total turbulent energy flux. Figure 2b shows a range in the simulated latent heat flux of about 75 $W m^{-2}$ in January (35 $W m^{-2}$ excluding $BUCK^C$) and about 100 $W m^{-2}$ in July. The main outlier was $BUCK^C$ which simulated anomalously high latent heat

fluxes during periods of significant rainfall. Note that GFDL[F] does not simulate anomalous sensible or latent heat fluxes due to the additional resistance term. UKMET[O] simulated high latent heat fluxes in August, and SPONSOR[L] simulated low latent heat exchange in May through to July. The range in the latent heat flux fell to about 40 $W m^{-2}$ for all months except July and August with these exceptions. Desborough (1997) shows that the parametrization of root distribution contributes to this scatter. A similar pattern was displayed by the models' simulation of the sensible heat flux (Fig. 2c).

3.2 Root zone soil moisture and runoff

The sixteen schemes simulated large differences in the root zone soil moisture (Fig. 2d) reached after a multi-annual equilibration period. However, the differences in the seasonal changes in soil moisture are much smaller than the differences in the annually averaged equilibrium soil moisture. There was no clear relationship between relative positions within the scatter of root zone soil moisture and evaporation. The range in simulated soil moisture remained similar through the year with CSIRO9[E], PLACE[J] and perhaps ISBA[H] higher than most and MOSAIC[I], BUCK[C] and SPONSOR[L] lower than most models. In the case of PLACE[J], the relatively high soil moisture is known to be related to the parametrization of the lower boundary condition and gravitational drainage. Some models which simulated relatively high annual total evaporation (BUCK[C], GISS[G]) simulated low soil moisture, while others (BASE[A], UGAMP[N]) simulated higher soil moisture levels. In contrast, SPONSOR[L] and SECHIBA2[K] simulated low annual evaporation and low soil moisture, while CSIRO9[E] and PLACE[J], despite high soil moisture amounts, simulated annually averaged evaporation in the centre of the overall range. Soil moisture in these schemes is effectively a simulated moisture index which is used to predict evaporation. Provided the functional relationship between an individual schemes' moisture index and evaporation is appropriate, differences in the amount of root zone moisture is not important to a GCM although this does hinder Improvement of schemes through difficulties in coupling to hydrological parametrizations and comparing to observations. The differences seen in Fig. 2d suggests that great care must be taken in exchanging soil moisture parametrizations between schemes unless the full relationship between moisture and evaporation is also exchanged. Overall, in this wet environment, the soil moisture to which a land surface model equilibrated did not appear to be related to the way it partitioned precipitation between evaporation and runoff.

Figure 2e shows that there are large differences in simulated total runoff (root zone plus surface). Two total runoff peaks are simulated, in February and October. The simulation of the February peak ranges from 575 mm $month^{-1}$ to 100 mm $month^{-1}$ (PLACE[J]). A similar range is simulated in October. CLASS[D] simulates the earlier runoff peak a month later than other schemes. From May to August, low precipitation leads to low runoff (mainly gravitational drainage) with many models simulating negative root zone drainage (i.e. ground water recharge). During this period, different levels of parametrization included in the schemes leads to a failure of the models to agree on the sign of the moisture flux.

There is no clear relationship between the complexity of the runoff parametrization included in individual land surface models and the confidence which can be placed in the resulting simulation. Simple schemes (e.g. BUCKET[C] or GFDL[F]) produce results which fall within the range of hydrologically more focused schemes (e.g. VIC[P], CLASS[D], PLACE[J]). Unfortunately, the overall range among the models is too large to establish whether the more sophisticated model is an improvement.

3.3 Temperatures

Figure 2f shows the model simulations of the seasonal effective temperature (the surface radiative temperature). The simulated differences in the surface temperature ranged between about 1.5 K (January) to 3.5 K in July. BUCK[C] was marginally warm in JJA (due to the excess evaporation earlier in the year and higher sensible heat flux in summer) while BASE[A] and MOSAIC[I] were slightly cold in July and August, corresponding to high evaporation rate. The simulation of canopy temperature (not shown) was very similar to the simulated effective temperature. There was no obvious relationship between complexity (in terms of number of soil layers, inclusion of a canopy for interception or inclusion of an explicit canopy model) and the simulation, nor did the seasonal effective temperature appear to depend on whether a full canopy temperature calculation was included. Therefore, in terms of the seasonal temperature cycle, the simplest schemes which have one layer and no canopy (e.g. BUCK[C], GFDL[F]) could not be clearly differentiated from the intermediate schemes with more soil layers and some canopy processes such as interception (e.g. UKMET[O]) or from the most complex schemes including multi-layer soil schemes and a reasonably complete canopy parametrization (e.g. BASE[A], CLASS[D] or MOSAIC[I]).

4 Seasonal results for the grassland simulation

The grassland forcing provided from the GCM included less rainfall and a more prolonged dry period

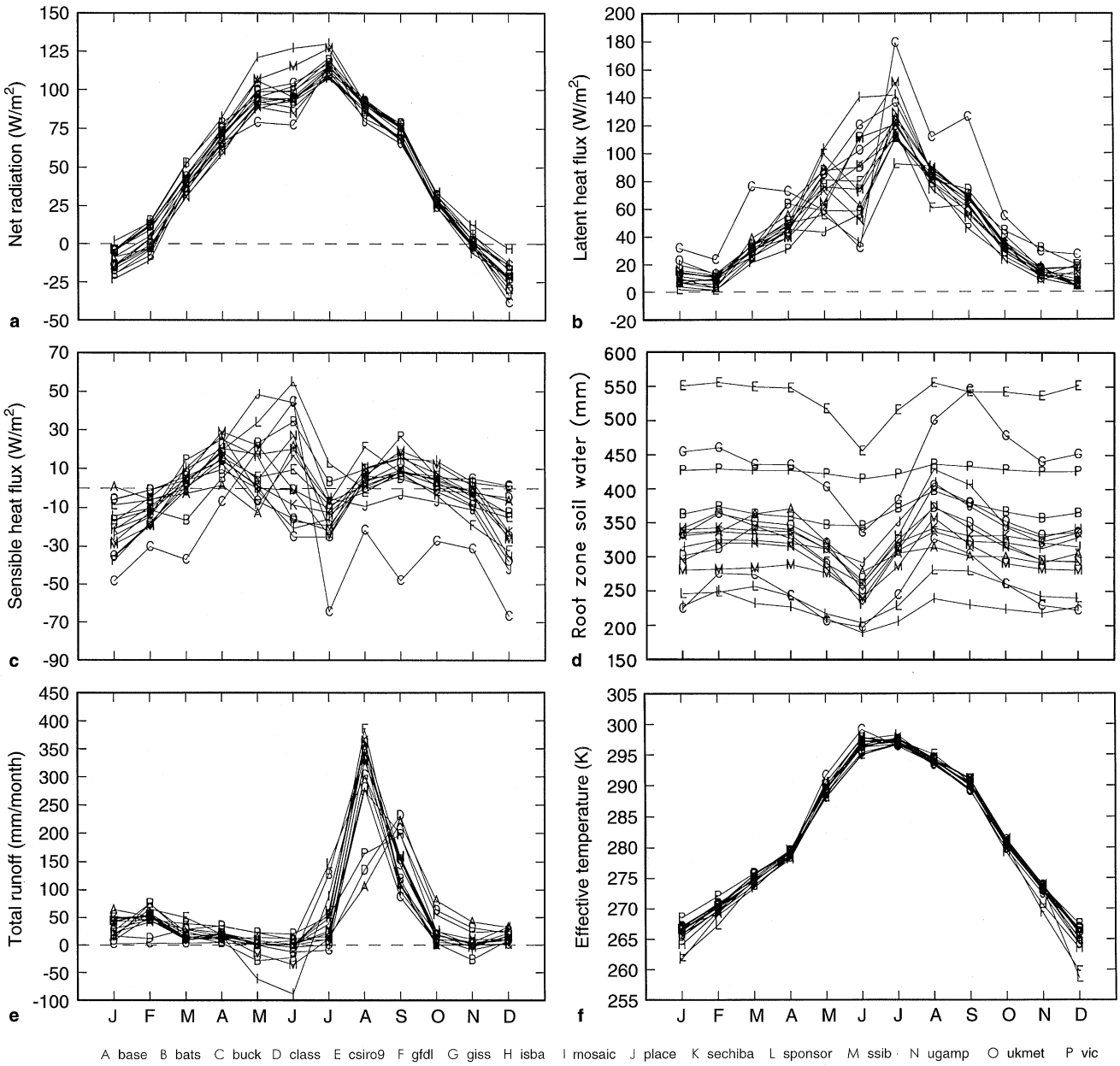

Fig. 3 a Mean monthly net radiation for GRASS; **b** as **a** but for the latent heat flux; **c** as **a** but for the sensible heat flux; **d** as **a** but for root zone soil moisture; **e** as **a** but for total runoff; **f** as **a** effective surface temperature

compared to the FOREST forcing. Precipitation occuring when the air temperature was below 0 °C fell as snow (snow fall occurred from November to February). This meant that snow accumulation and snow melt had to be simulated by each land surface scheme.

4.1 Energy fluxes

Figure 3a shows the net radiation for the GRASS simulation. As with FOREST, those models which used the Sellers et al. (1986) parametrization for canopy albedo tended to predict a lower snow-free albedo (not shown) than other models (largely due to difficulties in the choice of parameter values by PILPS). Under snow-free conditions, MOSAIC[I], and SSIB[M] all predicted low snow-free albedos relative to the other schemes hence the absorbed solar radiation was 15–20 W m^{-2} higher in these models in snow free conditions in early summer (not shown). PLACE[J] simulated a relatively low albedo in July but an anomalously high albedo in August and September, but this did not impact significantly on net radiation. Figure 3a also shows that BUCK[C] simulated anomalously low net radiation in May and June. This was caused by temperature rather than albedo differences. Overall, the

range in net radiation was about 30 $W m^{-2}$ in January. This increased to 45 $W m^{-2}$ in June (20 $W m^{-2}$ if MOSAIC[I] and SSIB[M] are omitted) and fell to only 15 $W m^{-2}$ in October. Most of these variations were due to albedo changes.

The problems of albedo in snow free conditions (where, because of the prescribed parameter values, the models simulated a range of about 0.1) was dwarfed by the differences in the albedos predicted by the land surface models during snow cover. MOSAIC[I], GISS[G], CLASS[D] and SSIB[M] predicted an albedo of less than 0.3 while GFDL[F] simulated albedos in excess of 0.7. This variability in simulated albedo represents a range in excess of 0.4, but this range remained about 0.25 if GFDL[F] and SPONSOR[L] were omitted. This led to a difference in the January absorbed solar radiation of 30 $W m^{-2}$ and indicates that the parametrization of the relationship between snow depth, snow cover and the effect of snow masking was a major source of differences among simulations at this time.

All models except BUCK[C] predicted a similar seasonal cycle in both the latent and sensible heat fluxes. There agreement in the latent heat flux in winter shows a range of about 30 $W m^{-2}$ (Fig. 3b) which remained 20 $W m^{-2}$ if BUCK[C] was excluded. VIC[P] simulated a slightly low latent heat flux in March and April while BATS[B] simulated a slightly high latent heat flux in October and November. During the summer months, when net radiation increased to around 100 $W m^{-2}$, scatter was much higher with the full range in June reaching 120 $W m^{-2}$ but with a central cluster of about half this range. The high simulation by MOSAIC[I] was due to the differences in the simulated albedo, but the cause of the very low latent heat flux simulated by SPONSOR[L] was not obvious. The increase in the range of the simulated latent heat flux (June) corresponds to a drying of the soil (see Sect. 4.4). The role of soil moisture stress was investigated further by Koster and Milly (1997). The scatter in the sensible heat flux (Fig. 3c) showed a similar (but reversed) pattern. The magnitude of the scatter in latent and sensible heat shown in Fig. 3b, c was not constant through the year when expressed as a percentage of net radiation. The scatter in the latent heat flux in January approximates the range in the net radiation, but in July, the scatter increased to about twice the range in net radiation.

4.2 Root zone soil moisture and runoff

As with FOREST, there was little agreement in either the magnitude of the amount or the shape of the seasonal soil moisture cycle (Fig. 3d). CSIRO9[E] simulated anomalously high soil moisture, but simulated a similar seasonal cycle to the majority of the other models. VIC[P] simulated a similar amount of soil moisture to the majority of schemes, but simulated no seasonal cycle because the model reaches a higher soil moisture level at equilibrium due to lower evaporation. GISS[G], MOSAIC[I], BUCK[C] and SPONSOR[L] also appeared anomalous during specific periods, but as with FOREST there was no consistent relationship with evaporation.

The models simulate a relatively similar seasonal variability in total (root zone plus surface) runoff (Fig. 3e). There are two periods where models are apparently anomalous. The first is in May and June when MOSAIC[I], BATS[B] and SSIB[M] simulate significant root zone recharge (this process is not parametrized by all models). The second period is in August and September when CLASS[D] BASE[A] and VIC[P] appeared anomalous simulating peak runoff a month later than the peak in rainfall.

4.3 Temperatures

Figure 3f shows that most models simulated the effective monthly average temperature to within a range of 5 K which is a lot higher than for FOREST. SPONSOR[L] is an exception and was about 6 K colder than most models during periods of snow cover. BUCK[C] was slightly warmer than most models in May and June. GFDL[F] was anomalously cold during periods of high snow cover. The anomalous simulations by SPONSOR[L] and GFDL[F] were due to the very high albedo simulated.

5 Discussion and conclusions

PILPS phase 1c aimed to identify the level of agreement between land surface schemes decoupled from atmospheric models (i.e. without the inclusion of surface-atmospheric feedbacks). We find that there is relatively poor agreement in the simulation of temperature, latent and sensible heat flux and runoff at the time scales analyzed.

There are many possible reasons for the differences discussed here. While every attempt was made to make the characterization of the surface as similar as possible, differences inevitably occurred due to the inconsistent use of terminology among land surface modellers, poor specification of some parameters, parameters having different effective values between schemes and the inability of some land surface modellers to capture some desired characteristics (e.g. albedo) as requested in their models. Part of the scatter may also be due to the lack of surface-atmospheric feedbacks in that a scheme can simulate an equilibrium climate which is inconsistent with the forcing because the equilibrium climate does not lead to an alteration in the atmospheric forcing. Therefore,

a small bias in a scheme may become over-emphasized. Finally, the atmospheric forcing, provided by a GCM, may have contained significant biases, such as an artificially low relative humidity, which might have exacerbated evaporation at the expense of sensible heat.

While the overall scatter may be due to a variety of causes independent of the physical realism of the land surface scheme, the difference between the simulations at the seasonal time scale is worrying. Despite identical atmospheric forcing and similar prescriptions of surface characteristics, the sixteen land surface schemes produced significantly different simulations of net radiation, temperature, turbulent energy fluxes, soil moisture and runoff. This is best illustrated with reference to latent heat fluxes. In the case of FOREST, the range in evaporation is between 50 $W m^{-2}$ (January) and 100 $W m^{-2}$ (August). Only about 20% of this can be explained by differences in the net radiation. In August, the fraction of the scatter which cannot be explained in terms of net radiation equates to about 80 mm of evaporation over the month. About 50% of the July scatter in latent heat for the GRASS simulation can be explained by differences in the net radiation. We therefore conclude that current land surface schemes simulate very different latent heat fluxes even when the surface is characterized in similar ways and the atmospheric forcing is prescribed. However, there is a question over how similar these land surface schemes can be expected to be given the experimental design. The 10 $W m^{-2}$ at the annual time scale would be expected to increase such that a 20–30 $W m^{-2}$ scatter at seasonal time scales would probably be expected. The land surface schemes shown here exceed this level of scatter, probably because of different ways of parametrizing moisture stress (on evaporation) and runoff terms (see Koster and Milly 1997). While these differences should not be ignored, if a few models are excluded, the overall scatter is reduced at the annual (but not at the seasonal) time scale. While this scatter is large, this work has not addressed the question as to whether these differences are actually large enough to affect a GCM's simulation. Coupled experiments are planned to address this issue.

It is generally not possible, from these experiments, to categorize any models as particularly good or particularly bad. Most models are designed for particular applications in which they have been tested and proven to perform well. The scatter between the models takes the form of a central cluster plus different outliers for each field. Recognizing that many of the PILPS models have evolved from a few original ancestors some clustering may be expected and it would therefore be wrong to assume that those models which do not agree with 'most' of the other models are necessarily flawed.

This study has generally described rather than explained the scatter simulated by sixteen land surface schemes. It is not possible to identify specific causes for the scatter, but in general all differences would have to be due to one of three causes: (a) structural differences in the models, (b) different choices of parameters (which can overlap a because leaving out some process can often be equated to setting some parameter to zero); and (c) bugs. Cause (b) has been reduced somewhat but not eliminated by PILPS. Cause (c) has also been partly addressed through extra quality control for phase 1(c). At this point we assume that the scatter is the result of the complex interaction between all the components of the schemes (i.e. largely cause a), but there is clearly much to be done to understand the specific causes of the scatter. Koster and Milly (1997) have begun this process through an analysis of how evaporation and runoff components interact in each scheme and Desborough and Pitman (1998) have investigated how differences in the parametrization of the surface energy balance effect the range of simulations discussed here. More work of this kind is a priority in order to resolve the disparity in the simulations and understand the reasons behind the scatter. Extending this work into the coupled environment is also a priority.

Acknowledgements We thank Dr. P. C. D. Milly for the considerable assistance he has given to PILPS. We also thank K.A. Dunne of the USGS for her considerable help in file handling and error checking. We would like to thank the support of the Australian Research Council, NOAA, the University of Arizona and DEST/NGAC for support of the PILPS project.

References

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

Budyko MI (1956). The heat balance of the Earth's surface. Gidrometeoizdat. Leningrad 255pp (In Russian)

Chen TH and 41 authors (1997) Cabauw experimental results from the Project for Intercomparison of Land-surface Parametrization Schemes. J. Clim 10:1194–1215

de Vries DA (1958) Simultaneous transfer of heat and moisture in porous media. Trans Am Geophys Union 39:909–916

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

Desborough CE (1997) The impact of root-weighting on the response of transpiration to moisture stress in land surface schemes. Mon Weather Rev 125:1920–1930

Desborough CE, Pitman AJ (1998) The BASE land surface model. Global Planet Change 19:3–18

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

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 72pp

Ducoudré NI, Laval K, Perrier A (1993) SECHIBA a new set of parametrizations of the hydrologic exchanges at the land/atmosphere interface within the LMD atmospheric general circulation model. J Clim 6:248–273

Garratt JR (1992) The atmospheric boundary layer. Cambridge University Press, Cambridge, UK, 316pp

Gedney N (1995) Development of a land surface scheme and its application to the Sahel. Unpublished Ph D Thesis, University of Reading, UK, 200pp

Gregory D, Smith RNB (1994) Canopy surface and soil hydrology. Unified Model Documentation Paper 25. UK Meteorological Office London Road Bracknell Berks RG12 2SZ, UK

Henderson–Sellers A, Pitman AJ, Love PK, Irannejad P, Chen TH (1995) The Project for Intercomparison of Landsurface Parametrization Schemes (PILPS): phases 2 & 3. Bull Am Meteorol Soc 76 : 489–503

Hillel D (1982). Introduction to soil physics. Academic Press, New York, 364pp

Koster R, Milly PCD (1997) The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. J Clim 10 : 1578–1591

Koster R, Suarez M (1992) Modeling the land surface boundary in climate models as a composite of independent vegetation stands. J Geophys Res 97 : 2697–2715

Kowalczyk EA, Garratt JR, Krummell PB (1991) A soil-canopy scheme for use in a numerical model of the atmosphere–1D stand-alone model. CSIRO Division of Atmospheric Research: Techn Pap 23: 56pp

Liang X, Lettenmaier DP, Wood EF, Burges SJ (1994) A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J Geophys Res 99 : 14415–14428

Liang X, Wood EF, Lettenmaier DP (1996) Surface soil moisture parametrization of the VIC-2L model: evaluation and modifications. Global Planet Change 13 : 195–206

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

Milly PCD (1992) Potential evaporation and soil moisture in general circulation models. J Clim 5 : 209–226

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

Philip JR, de Vries DA (1957) Moisture movement in porous materials under temperature gradients. Trans Am Geophys Union 38 : 222–232

Pitman AJ and 41 authors (1993) Results from the off-line control simulations (phase 1a) of the Project for Intercomparison of Land surface Parametrisation Schemes (PILPS) GEWEX Techn Note IGPO Publi Ser 7 : 47pp

Robock A, Vinnikov KY, Schlosser CA, Speranskaya NA, Xue Y (1995) Use of midlatitude soil moisture and meteorological observations to validate soil moisture simulations with biosphere and bucket models. J Clim 8 : 15–35

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

Shmakin AB, Mikhailov AY, Bulanov SA (1993) Parametrization scheme of the land hydrology at different spatial scales In: Bolle H-J, Feddes RA, Kalma JD (eds) exchange processes at the land surface for a range of space and time scales. IAHS Publ 212 : 569–575

Verseghy DL (1991) CLASS: a Canadian land surface scheme for GCMs: I soil model. Int J Climatol 11 : 111–133

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

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 94pp

Wetzel PJ, Boone A (1995) A parametrization for land-atmosphere–cloud exchange (PLACE): Documentation and testing of a detailed process model of the partly cloudy boundary layer over heterogeneous land. J Clim 8 : 1810–1837

Wetzel PJ, Chang JT (1988) Evapotranspiration from nonuniform surfaces a first approach for short term numerical weather prediction. Mon Weather Rev 116 : 600–621

Xue Y, Sellers PJ, Kinter JL, Shukla J (1991) A simplified biosphere model for climate studies. J Clim 4 : 345–364