J 5.6

PILPS SEMI-ARID EXPERIMENT: PRELIMINARY RESULTS

Luis A. Bastidas*[1], Enrique Rosero[1], Bart Njissen[2]
[1]Civil and Environmental Engineering and Utah Water Research Laboratory, Utah State University, Logan, Utah
[2]Hydrology and Water Resources, University of Arizona, Tucson, Arizona

1. INTRODUCTION AND SCOPE

The Project for Intercomparison of Land-surface Parameterizations Schemes (PILPS) is part of the Global Land Atmosphere System Study (GLASS) whose goal is to improve the understanding of the parameterization of the interactions between the atmosphere and the continental surface *[Pitman and Henderson-Sellers, 1998]* though intercomparisons of current state-of-the-art parameterization schemes. GLASS is part of GEWEX, the Global Energy Water Experiment, which is a program of the World Climate Research Program (WCRP) *[Lettenmmaier, 2003]*.

The PILPS semi-arid experiment (also known as PILPS San Pedro) has unique characteristics. It not only focuses on a different environment from previous PILPS experiments, but it also will employ appropriate system methods for parameter estimation, that will help the modeling groups in identifying parameter sets that make the models consistent with the observational data.

In this preliminary work, two 4+ year long data sets from the USDA Experimental Watershed in the Walnut Gulch, Arizona *[Emmerich, 2003]*, are used for model intercomparison analysis with focus on evaluating the consistency the net energy partition into latent, sensible and ground heat flux components. Additionally, because the data sites correspond to two different semi-arid environments (shrub and grass) we attempt to diagnose the extent to which the standard representation of the eight participant models is adequate.

2. BACKGROUND

Despite the fact that 1/3 of the global land surface is semi-arid or arid, none of the previous PILPS experiments *[e.g., Njissen et al., 2003, Boone et al., 2001; Chen et al., 1997, Lettenmaier et al., 1996, Pitman et al.,1993, Henderson-Sellers et al.,1995, Henderson-Sellers et al.,1993]* has focused on comparing the performance of land-surface models in such environments.

In water controlled regions such as the arid U.S. Southwest, vegetation is highly dependent on imposed stresses. Its response to precipitation is distinct and variable, i.e. grass species (C_4) respond quickly to upper soil moisture during the summer monsoon while shrub plants (C_3) tend to use deeper soil moisture reserve and being active in spring and fall *[Scott et al., 2000, Kemp, 1983]*. The role played by the biomes as a linkage between soil moisture and evaporation process is especially relevant since soil moisture strongly controls the nature of water, energy and momentum fluxes, such as the partitioning of available energy between latent and sensible heat *[Entekhabi and Rodriguez-Iturbe, 1994]*.

Predicting the availability of water resources in these hydrologically stressed regions depends fundamentally on the ability to understand and reproduce the interaction of vegetation processes with climate and its effects on the water cycle. However, many models have been working under the assumption that semi-arid areas are homogeneous (therefore, assigning similar set of parameters for all biomes) and even considering them as bare-soil areas neglecting the interaction vegetation-atmosphere *[Bastidas et al., 2001, Bastidas et al., 2002]*.

Accurate representation within the modeling framework used for land-surface-atmosphere schemes is crucial not only for analyzing the current state of the system but also for making predictions of potential climate change impacts, assessing the effect of changes in vegetation type (i.e., shrub invasion) or increase in demand due to population growth *[Hogue et al., 2005].*

The unique characteristics of PILPS San Pedro allows not only to asses the ability of the models to reproduce the water and energy exchanges in semi-arid environments but also to test if the current (usually single) parametric representations of semi-arid lands in the LSM are enough to simulate the different environments *[Bastidas et al., 2003]*.

3. SITES, MODELS AND DATA

3.1. Lucky Hills and Kendall Sites

The initial part of the experiment has been carried out at two sites within the Walnut Gulch Experimental Watershed in Southeastern Arizona, a sub-basin of the Upper San Pedro Basin.

The Lucky Hills site (110°03'05'' W, 31°44'37'' N) is located in the lower (1372 masl) shrub dominated part of the basin. The vegetation consists mainly of the C_3 species *[Scott et al., 2000]*. Soils are mostly loamy sand or very gravelly sandy loams. Canopy height is estimated at 1 m. Slopes are 3-8%. Average temperature is 18.6 °C.

The Kendall site (109°56'28'' W, 31°44'10'' N) is in the eastern part of the watershed covered mainly by perennial C_4 grasses. The elevation is 1526 masl. Soils consist mainly of very gravelly sandy loams which contain limestone rock fragments. Canopy height is estimated 0.4 -0.7 m. Slopes are 4-9%. Average temperature is 19.3 °C.

* *Corresponding author address*: Luis A. Bastidas, Utah Water Research Laboratory, Utah State University, Logan, UT 84322-8200; e-mail: luis.bastidas@usu.edu

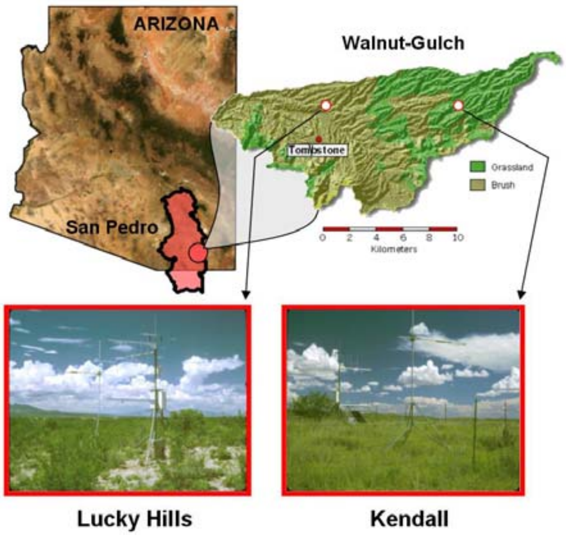

Figure 1. Lucky Hills desert shrub and Kendall grassland sites within the USDA Walnut Gulch Experimental Watershed.

The annual precipitation in the U.S. Southwest is bimodal. About 2/3 of the yearly amount of precipitation occurs under the influence of the North American Monsoon (NAM), where events have high intensity and short duration. Soil moisture recharge is insignificant since plant transpiration and bare soil evaporation are high. During winter, frontal precipitation events have longer duration and less intensity allowing recharge at lower depth.

Table 1. 1964-1994 Average annual, monsoon and winter precipitation in mm for the Lucky Hills and Kendall Sites *[Scott et al., 2000]*.

Site	Annual	Jul-Sep	Nov-Feb
Lucky Hills	338	200	80
Kendall	351	200	86

3.2. Data provided

Precipitation and weather data (net radiation, ground heat flux, wind speed, wind direction, air temperature, and relative humidity) have been collected continuously by the USDA-ARS Tucson from January 1997 to December 2000 at 20-minute intervals using a Bowen ratio system with a tower height of 3 m *[Emmerich et al., 2003].* It includes measurements of sensible and latent heat fluxes, CO_2 flux and soil temperature.

3.3. Participant Models

A. BATS, UCLA, USA
B. CBM, CSIRO, Australia
C. ISBA, Meteo France
D. Noah, USU/NCEP, USA
E. SEWAB, GKSS Research Center, Germany
F. SiB 2, USU – not reported here.
G. SPONSOR, Institute of Geography, Russian Academy of Sciences
H. SSiB, COLA, USA
I. SWAP, Institute of Water Problems, Russian Academy of Sciences

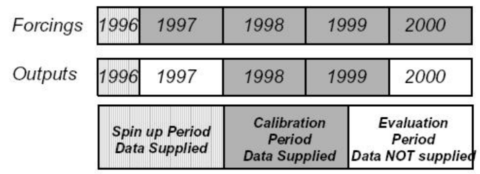

Figure 2. Data supplied to participants in dark gray. *[Bastidas et al., 2003]*

4. METHODS

Four+ years of quality controlled forcing data and a subset of two consecutive years of evaluation data were provided to the participants (Figure 2).

Phase 1 of the experiment requested off-line runs using the provided forcing and default parameters (a priori established model parameters for semi-arid regions) that included several energy and water balance model outputs and state variables following the ALMA standards.

Split sample validation tests were performed on those results using goodness of fit measures (R^2, IA, NSE, RMSE, Bias), similarity measures (Hausdorff Norm), statistical correlation (Taylor diagrams), etc.

The considered time scales for the analysis were: monthly, daily average, annual, interannual, and seasonal with especial attention to the NAM.

5. ANALYSIS RESULTS AND DISCUSSION

5.1. Energy balance closure

The model evaluation was carried out by comparing their outputs against long-term measurements of surface sensible and latent heat fluxes, ground and total net radiative fluxes. Temporal aggregations were typically made to evaluate model performance at yearly time scales *[Wood et al., 1998, Boone et al., 2001]*.

The interannual energy balance check is presented on Figure 3b. The abscissa gives the mean annual ground heat flux in wm^{-2} for the whole period and for each year. The ordinate shows the mean annual energy residuals over the same period accordingly. If energy is conserved, the annual energy residual, defined as $\varepsilon=R_n-Q_{le}-Q_h-Q_g$, should be zero. All of the schemes have residuals less than ± 3 wm^{-2} which has been the criterion used in previous PILPS experiments *[Chen et al, 1997, Wood et al, 1998]*. In fact, it can be seen that only model G (SPONSOR) has a residual larger than 2 wm^{-2} (at Kendall site -1998). Model D (Noah) is every year slightly on the negative side.

The mean ground heat flux is also expected to be close to zero. This is true for models C (ISBA), D

(Noah), H (SSIB) and I (SWAP). Model A (BATS) loses ground heat flux in both sites, up to -4.4 wm^{-2} in Lucky Hills and -3.58 in Kendall. Model E (SEWAB) shows negative Qg only in Lucky Hills. On the other hand model G (SPONSOR) has the highest annual mean ground flux of +6.17 wm^{-2} in year 1998.

It should be noticed that in Lucky Hills the observed interannual mean ground heat flux is slightly negative in every period (-0.43 wm^{-2} on average). In Kendall, it is very close to zero since alternates between a year negative and the next positive.

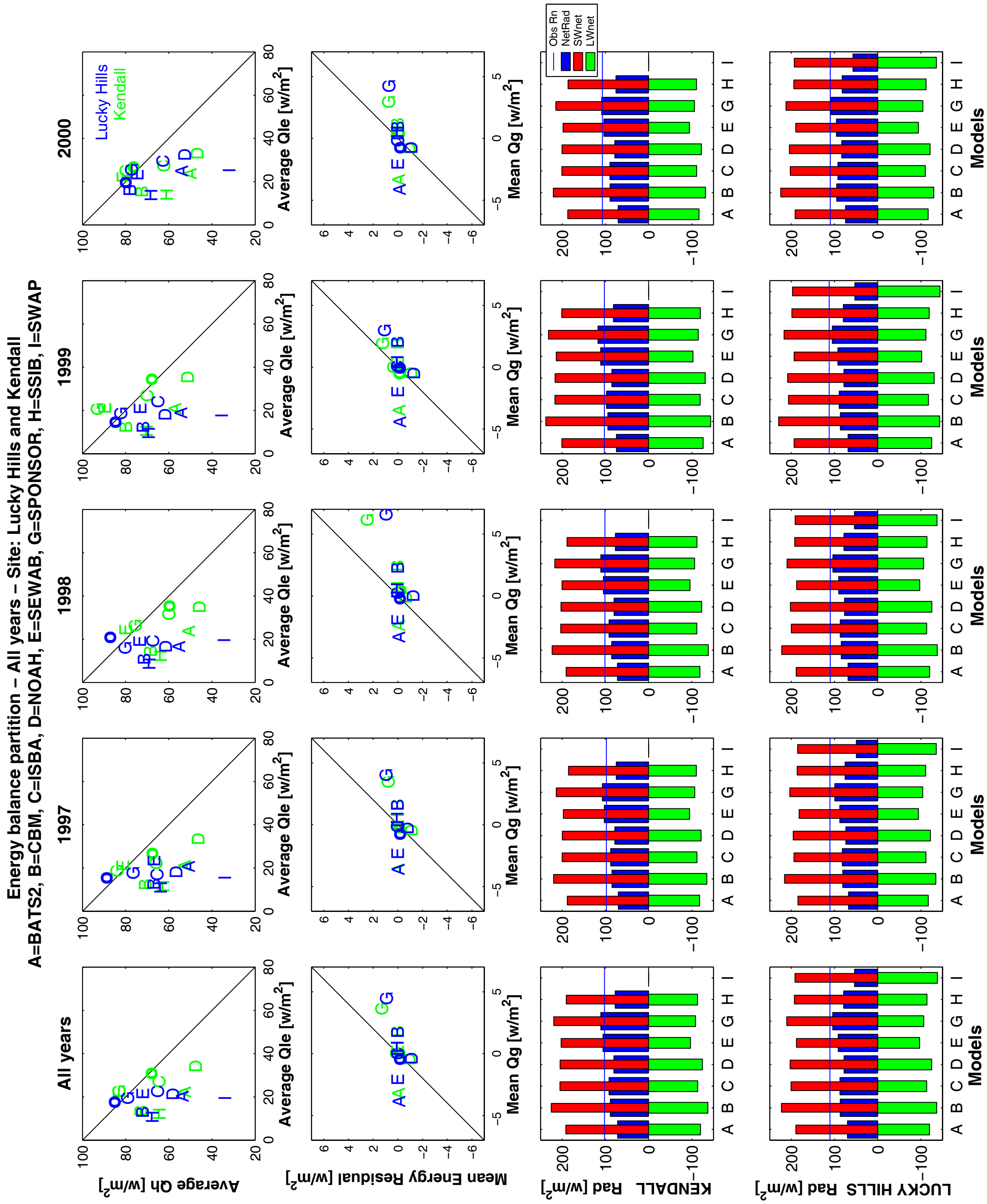

Figure 3. Energy Balance closure and net radiation decomposition for both sites and all periods. (a) Energy partition: Sensible heat vs. Latent heat (b) Energy balance closure: Mean energy residual vs. ground heat flux. Long and shortwave components of Net Radiation in (c)Kendall and (d)Lucky Hills.

Assuming energy conserving schemes for all models, the analysis can focus on the mean annual sensible and latent heat components of the energy balance compared to the net radiation.

5.2. Energy balance decomposition

The net energy partition is presented in Figure 3a, where the mean annual sensible heat is plotted against the mean annual latent heat. It is expected that models map along the line of net radiation. Observed mean annual net radiation is in the order of 100 wm^{-2} at Kendall and 10% larger at Lucky Hills. For simplicity only one line is plotted, corresponding to 100 wm^{-2}.

The observation-derived partition Q_h:Q_{le} for all years is 68.8 : 30.8 in Kendall (green O) and 86 : 17.7 in Lucky Hills (blue O). In general, all the models report larger sensible heat flux than latent heat; however they do not lie in the line suggesting an underestimation in the computed net radiation. Accordingly, models who report low latent heat also present relative warmer surfaces and lower net radiation. Most models predict a soil temperature between 288.9 and 298.9 K in Lucky Hills and 288.1 and 297.6 K in Kendall when the observed is 293.3 K and 292.6 K respectively. The among-models range of about 10 K is the same order that the observed standard deviation.

There is a considerable scatter in sensible heat (Q_h) values. Model I (SWAP) (Q_h=33.9 wm^{-2}) is an outlier that has the smallest energy available to partition.

All models underestimate Q_h for Lucky Hills, the among-models range is 53-79 w/m^{-2}. In Kendall, model outputs are within 20% of the observed value. The energy partition showed by models E (SEWAB) and G (SPONSOR) is very similar at Kendall for the entire period, however some differences are apparent on a year by year basis; H (SSIB) and B (CBM) show similar partition in both sites for all the years. The Kendall site has more scattered latent heat (Q_{le}) values than Lucky Hills except for the year 2000 when they are almost the same.

The mean annual Bowen ratio B=Q_h/Q_{le} varies between sites and models. Table 2 presents the Bowen ratios of the schemes based on the interannual mean annual sensible and latent heat, in comparison to the observation-derived one. As reported in the literature for semi-arid areas *[Nobel, 1999]*, Bowen ratios in the range 2 to 6 have been computed by the models. The difference in the observations suggests that different schemes predict different energy balance climatologies and that depending on the model probably different physical mechanism may be controlling the exchanges at the surface. Only model D (Noah) has the same relationship between sites, as observed (~1:2). For the rest of the models, the ratio between sites is ~4:5 except for model E (SEWAB) where Kendall's Bowen ratio is larger than Lucky Hills'. Annual mean latent heat flux in Kendall is almost 75% larger than in Lucky Hills. Four models (A,B,E,H) present the same partitioning independently of the site. This means that the models are not making appropriate distinctions between the dominant biomes in the sites.

The observed behavior, i.e., models not mapping along the net radiation line, is explained in Figure 3 c,d. It shows the disaggregation of net radiation (blue) into short (red) and longwave (green) components, for all models and in all periods. Observed net radiation is represented with the blue line across the bars. Computed mean annual net radiation varied between 71.6 and 110.7 wm^{-2} in Kendall and between 53.43 and 103.94 wm^{-2} in Lucky Hills. All models underestimate the net radiation in Lucky Hills. Model I (SWAP) is an outlier whose net radiation is less than 50% of the observed. The remaining models report net radiation values of up to 70% of observed. In Kendall models E (SEWAB) and G (SPONSOR) have net radiation values larger than observed. The rest of the models have values of net radiation up to 75% of the observed. Thus, it can be said that the intermodel differences are similar for the long and shortwave radiation components.

Table 2. Interannual Bowen ratio for both sites and all models.

	Kendall	**Lucky Hills**
Observed	2.23	4.86
A	2.34	2.65
B	5.35	5.57
C	2.39	2.89
D	1.38	2.71
E	3.74	3.33
F	3.63	4.02
G	5.3	6.27
H	-	1.73

In Figures 4 and 5 we present the Taylor diagrams for the two sites. They summarize the model ability to reproduce the observations. Taylor diagrams allow representation of three important (second-order) statistics of the performance in one plot. The distance from the origin to a point is the standard deviation and the azimuth angle is the arccosine of the correlation between observed and computed. The observed value is plotted along the horizontal axis. The distance from there to any point corresponds to the root mean squared error (RMSE). Simulated Net Radiation (NetRad) and sensible heat (Q_h) have low RMSE and high correlation. Mean latent heat (Q_{le}) present low correlation r=0.5. More differences among models are found in the representation of ground heat flux (Q_g).

In Figure 6, for the Lucky Hills site, we present the normalized (reduced variable) mean monthly model simulated energy balance components for the participating models and the observations. The computed values for the similarity measure (Hausdorff norm) are also presented – the lower the value the better the match. The similarity measure allows for the simultaneous evaluation of the performance of the four components of the energy balance. Figure 7 depicts the same for Kendall. Models G (SPONSOR) and C (ISBA) perform better in Kendall while E (SEWAB) does so in Lucky Hills. CBM, Noah, and SWAP have troubles reproducing the monthly average values at the Kendall site, particularly the Q_{le}. The problems are not so apparent at the Lucky Hills site. It is of interest to note that all the models perform better at the Lucky Hills site than at the Kendall site. An explanation could be that the models are geared towards representing shrubs than grasses in the semi-arid regions.

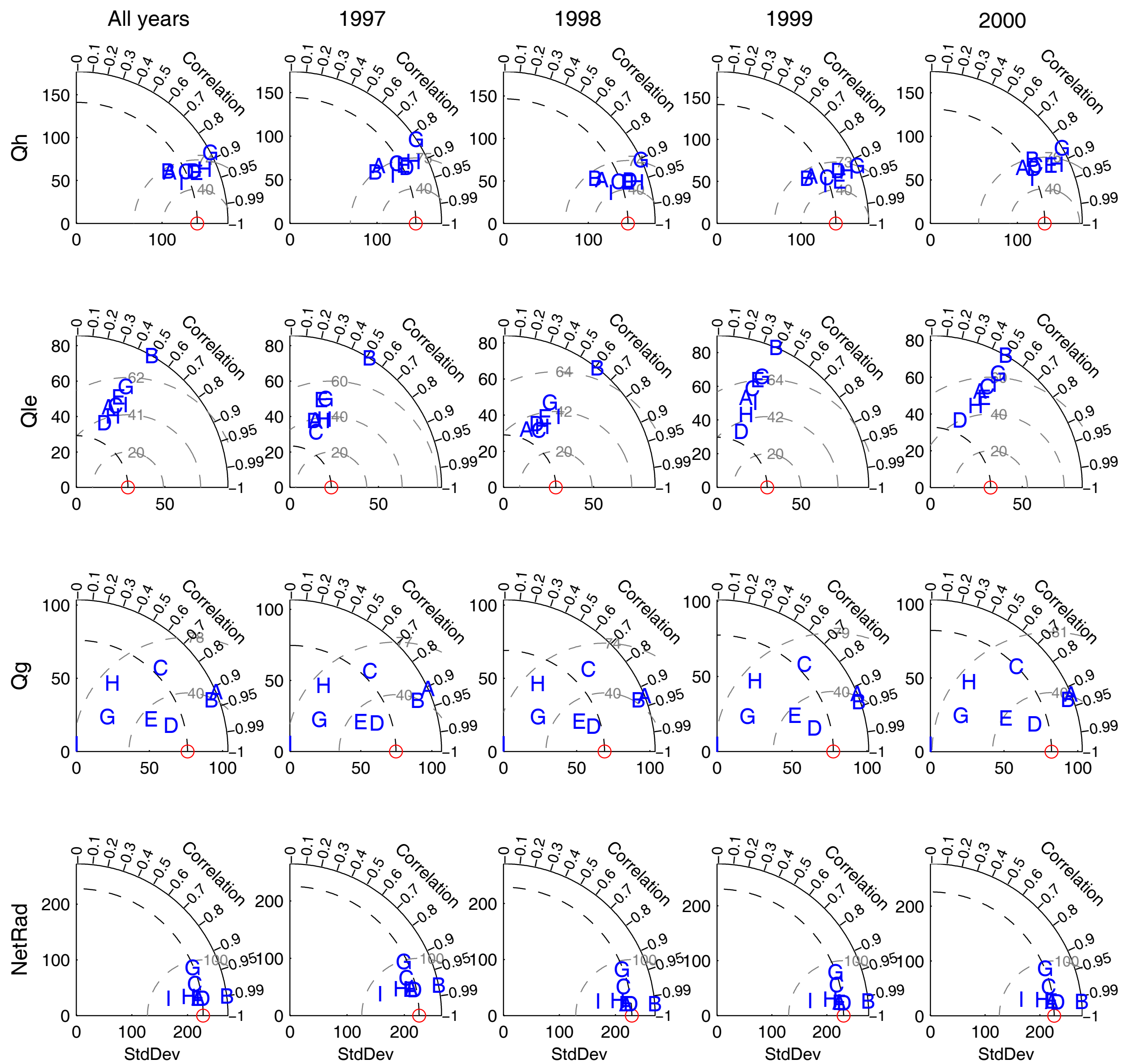

Figure 4. Taylor diagram of energy related variables in Lucky Hills

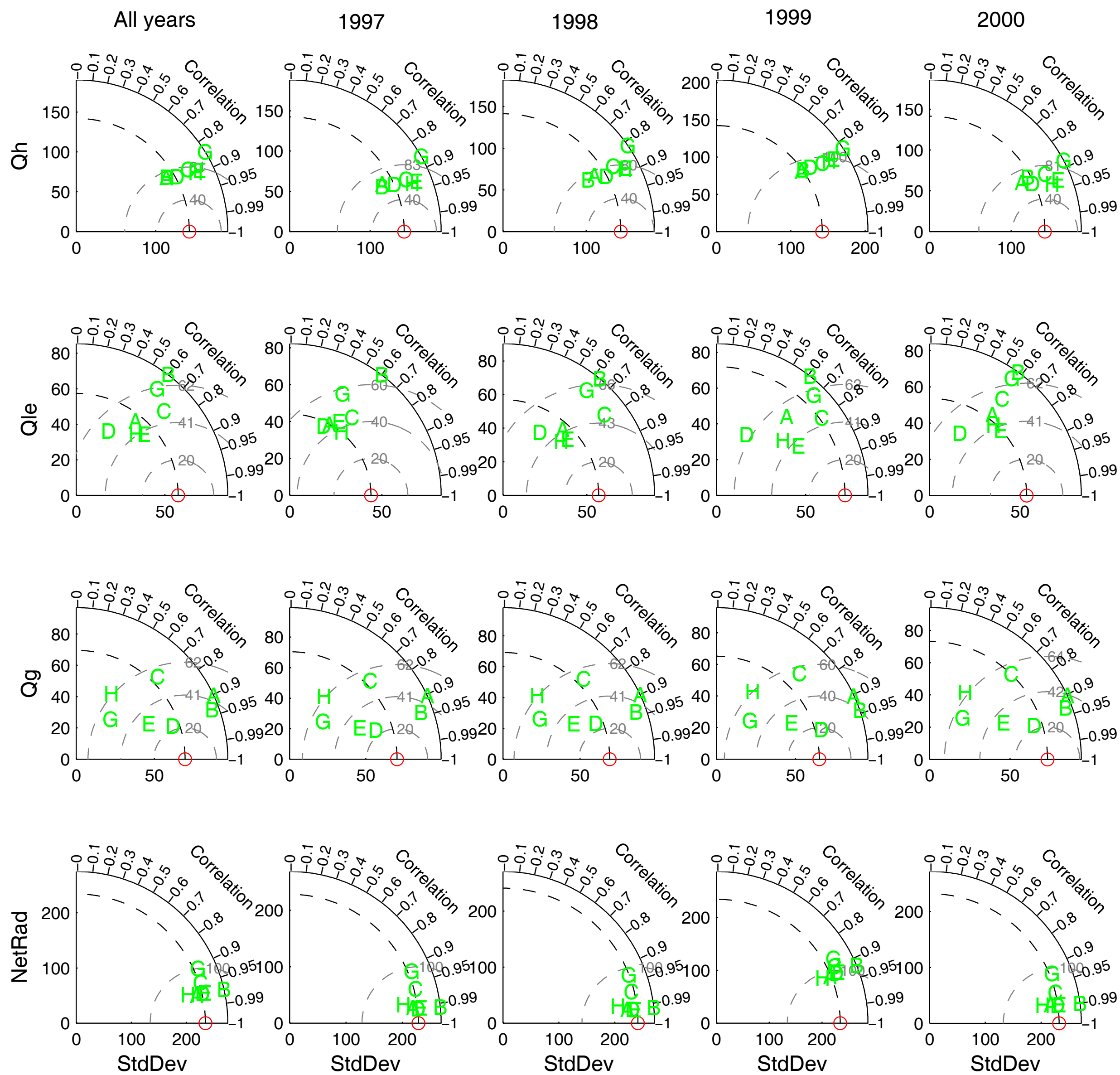

Figure 5. Taylor diagrams of energy related variables in Kendall.

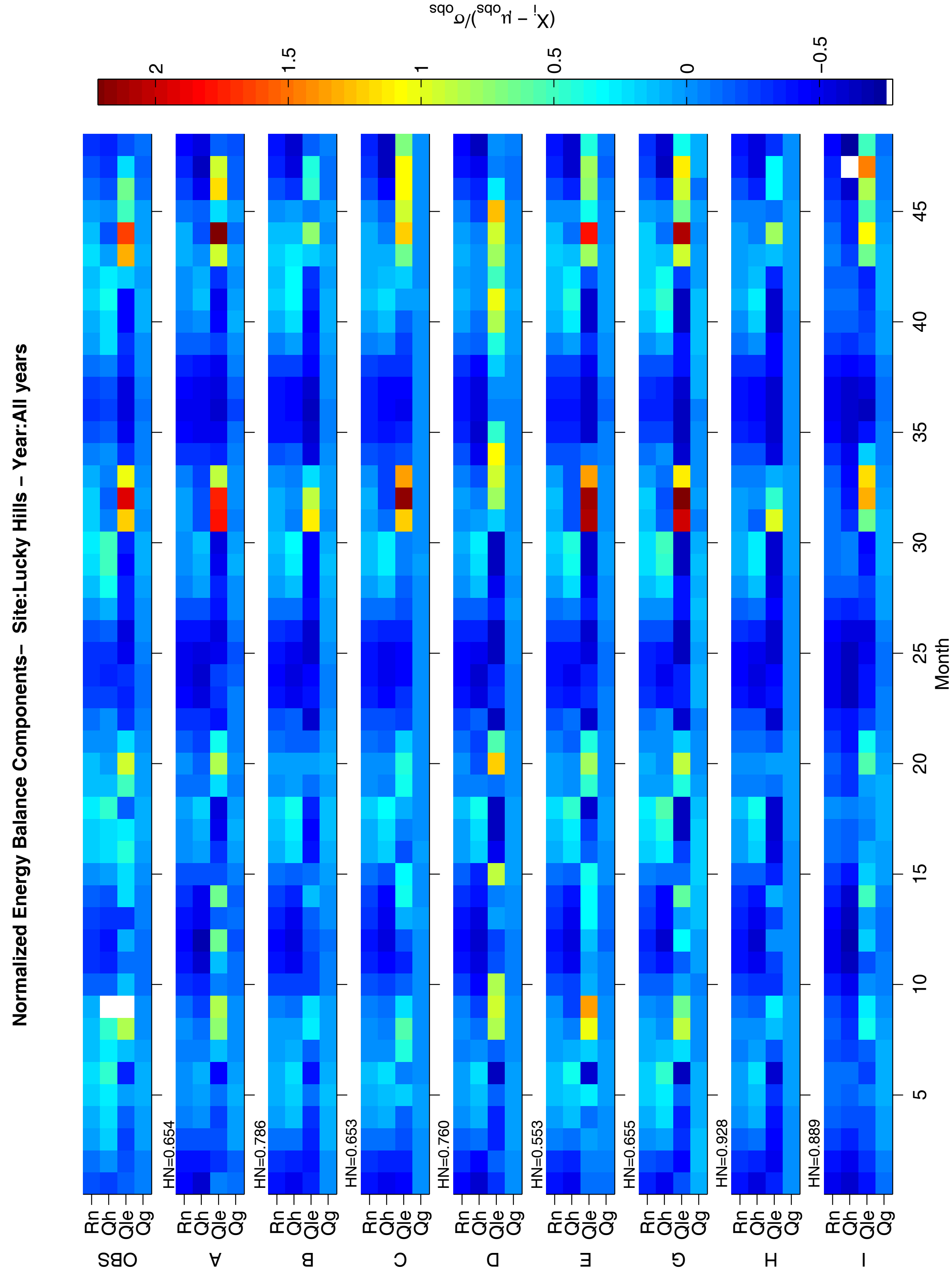

Figure 6. Similarity based comparison between observed mean monthly reduced variables and model representations of energy balance components in Lucky Hills

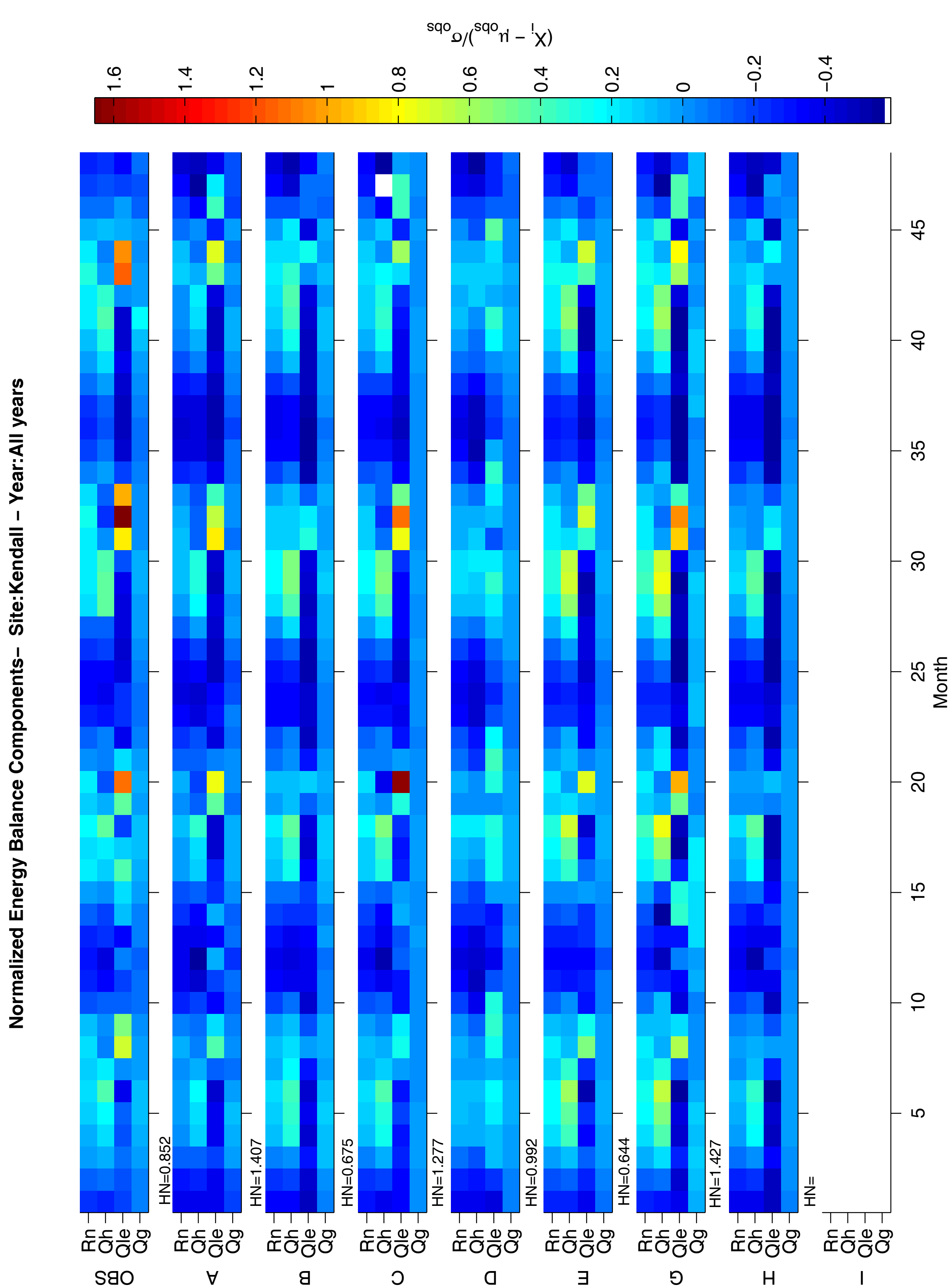

Figure 7. Similarity based comparison between observed mean monthly reduced variables and model representation of energy balance components in Kendall

The results are consistent with the initial hypothesis that a general representation of semi-arid areas needs revision. Semi-arid regions constitute a complex environment within which an easy transferability of characteristics is not appropriate.

6. SUMMARY AND CONCLUSIONS

Interannual closure check showed that all of the schemes have energy balance residuals less than ± 2 wm^{-2}. The mean annual ground heat flux is zero for three models (C,H,I). Three models (A,E,D) lose ground heat flux and two (B,G) have positive mean annual ground heat fluxes. Models have failed to adequate represent the annual net radiation. All of the models underestimate on the net radiation in Lucky Hills and only two (E,G) are not below the observed net radiation in Kendall. One model (I) is an outlier whose computed net radiation is less than 50% of observed. Large differences in the surface energy partitioning between latent and sensible heat fluxes have been observed among models and between sites. Bowen ratio for Lucky Hills site (shrub) is 4.86 opposed to 2.23 for the Kendall site (grass). The among-model range of Bowen ratios for the grassland site varies from 1.38 to 5.30. In the shrub site it ranges between 2.65 and 6.27. The differences found in Bowen ratio values suggest the possibility that the different biomes are not being represented adequately. In Lucky Hills most of the models (except H,B) represent well the annual latent heat flux but underestimate on the sensible by more than 15%. In Kendall, only one model (D) represents the mean annual latent heat adequately and the rest of them underestimate it but three (B, E, G) overestimate on the sensible heat and the underestimation of the others is not larger than 20%.

Because, at this stage of the experiment participants have not provided a calibrated output, it is difficult to conclude if the aforementioned shortcomings are to be ascribed mainly to model structural defficiencies or if the default parameter sets are not the appropriate ones for the extreme heterogeneity present in semi-arid environments. In the next stages of the PILPS San Pedro experiment a multi-objective parameter estimation approach using optimization algorithms will be applied to improve the performance of LSMs by constraining the parameters to make the outputs consistent with the observational data *[e.g. Bastidas et al., 2005, 2002, 2001, 1999, Xia et al., 2002, Gupta et al., 1999]*. Previous work by several authors showed that by means of calibration, not only the divergence between observations and model outputs has been minimized but also frequently, more (physically) meaningful parameter values have been found. In this context it is desirable to have calibrated runs to evaluate the performance of the models under similar (and more fair) conditions, i.e. the parameter error is removed leaving only model structural and data errors as the only components of the overall simulation error.

The next stages of the PILPS San Pedro Experiment will also address the issue of parameter transferability between sites with different vegetation covers, located a few kilometers apart and the transferability of parameters between sites with similar conditions but hundreds of kilometers apart. These are some of the characteristics that make the PILPS San Pedro Experiment an interesting one for modeling groups to participate in.

7. ACKNOLEDGEMENTS

Primary support for this study was provided by the Utah Water Research Laboratory and SAHRA (Sustainability of Semi-Arid Hydrology and Riparian Areas) under the STC Program of the National Science Foundation Agreement N° EAR-9876800.

8. REFERENCES

Bastidas, L.A., B. Nijssen, W. Emmerich, H.V. Gupta and E. Small, 2003, Description of the PILPS 2g Experiment: Model Comparison over Semi-Arid Areas. Proposal submitted to the GLASS science panel of GEWEX.

Bastidas, L.A., H.V. Gupta, and S. Sorooshian, 2001. Bounding parameters of land surface schemes with observational data, in Observations and Modeling of the Land Surface Hydrological Processes, V. Lakshmi, J. Albertson & J. Schaake, Editors, Water Science and Application Volume 3, p.65-76, AGU.

Bastidas, L.A., H.V. Gupta, and S. Sorooshian, 2002. Parameter, Structure and Performance Evaluation for Land Surface Models, in Advances in the Calibration of Watershed Models, Q. Duan, H.V. Gupta, S. Sorooshian, A. Rousseau and R. Turcotte, Editors, AGU.

Boone, A., F. Habets, and J. Noilhan, 2001: The Rhone-AGGregation Experiment. *GEWEX News*, WCRP, **11**(3), 3-5.

Chen, T. H., A. Henderson-Sellers, P. C. D. Milly, A. J. Pitman, et al., 1997: Cabauw experimental results from the project for intercomparison of landsurface schemes (PILPS). *J. Climate*, **10**, 1194-1215.

Emmerich, W.E., 2002. Carbon dioxide fluxes in a semiarid environment with high carbonate soils, Agri*cultural and Forest Meteorology,* 3105, 1-12.

Entekhabi, D. and I. Rodriguez-Iturbe, 1994. Analytical framework for the charchterization of the space-time variability of soil moisture, *Advances in Water Resources*, **17**, 35-45.

Gupta, H.V., L.A. Bastidas, S. Sorooshian, W.J. Shuttleworth and Z.L. Yang, 1999, Parameter estimation of a Land Surface Écheme using Multi-Criteria Methods, *Journal of Geophysical Research*, Vol. 104, NoD16, p.19,491-19,504.

Henderson-Sellers A., Pitman A., Love P., Irannjead P., Chen T., (1995) The Project for Intercomparison of Land-Surface Parametrization Schemes (PILPS): Phases 2 and 3, *Bulletin of the American Meteorological Society,* **76**: 489-503

Henderson-Sellers A., Yang Z., Dickinson R., (1993) The Project for Intercomparison of Land-Surface Parametrization Schemes, *Bulletin of the American Meteorological Society,* **74**: 1335-1349

Hogue, T.S., L.A. Bastidas, H. Gupta, S. Sorooshian, and K. Mitchell, Evaluation and Transferability of the Noah Land Land Surface Model in Semiarid Environments, *Journal of Hydrometeorology*, V 6, pp. 68-84, 2005.

Lettenmaier, D.P., 2003, Preface, *Global and Planetary Change*, 38, pp. vii-ix.

Lettenmaier, D., D. Lohmann, E.F. Wood, and X. Liang, *PILPS-2c Workshop Report*, Princeton Univ., Princeton, N.J., October 8-31, 1996.

Njissen, B., L.C. Bowling, D.P. Lettenmaier, et al., 2003. Simulation of high latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e) 2: Comparison of model results with observation, *Global and planetary Change*, **38**, 31-53.

Nobel., P.S., 1999, Physicochemical and environmental plant physiology. San Diego: Academic Press, 474 p.

Pitman, A. and A. Henderson-Sellers, Recent progress and results from the project for the intercomparison of landsurface parameterization schemes, *Journal of Hydrology*, 212-213, pp. 128-135.

Pitman, A., Henderson-Sellers A., Abramopoulos F., Avissar R., Garratt J., et al., 1993, Project for Intercomparison of Land-Surface Parametrization Schemes (PILPS). Results from off-line control simulations (Phase 1A). Gewex report IPGO Publication.

Scott, R.L., W. Shuttleworth, T. Keefer. 2000. Modeling multiyear observations of soil moisture recharge in the semiarid American Southwest, *Water Resources Research*, 36, 8, 2233-2247.

Wood, E.F., D. Lettenmaier, X. Liang, 1998. The project for Intercomparison PILPS 2(c) Red J. Global and Planetary Change, 19

Xia, Y., A.J. Pitman, H.V. Gupta, M. Lepastrier, A. Henderson-Sellers and L.A. Bastidas. 2002. Calibrating a land surface model of varying complexity using multi-criteria methods and the Cabauw data set, *Journal of Hydrometereology*, V3(2), pp.181-194