Description of the PILPS San Pedro-Sevilleta Experiment Model Comparison over Semi-Arid Areas

Proposed by

Luis A. Bastidas[1,2]
Hoshin V. Gupta[2,3]
Bart Nijssen[2,3,4]
William Emmerich[5]
Eric Small[6]

to the GLASS Science Panel

Revison 1: October 2002
Revision 2: December 2002
Revision 3: June 2003
Revision 4: January 2004
Revision 5: February 2004

[1] Department of Civil & Environmental Engineering and Utah Water Research Laboratory, Utah State University, Logan, Utah
[2] SAHRA – NSF Science and Technology Center for Sustainability of Semi-arid Hydrology and Riparian Areas, University of Arizona, Tucson, Arizona
[3] Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona
[4] Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, Arizona
[5] U.S. Department of Agriculture, Agricultural Research Service, Tucson, Arizona
[6] Department of Geosciences, University of Colorado, Boulder, Colorado

CONTENT

1. INTRODUCTION

1.1 *Objective of PILPS San Pedro-Sevilleta*

The PILPS experiments conducted so far have been important for the development and evaluation of land surface models *[Pitman et al, 1993, Henderson-Sellers et al., 1993, Henderson-Sellers et al., 1995]*. However, these experiments have not included any on semi-arid lands, despite the fact that 1/3 of the global land surface is semi-arid or arid. It is imperative, thus, to carry out a PILPS experiment over semi-arid lands.

The PILPS San Pedro-Sevilleta experiment proposed here is an initiative within the GEWEX/GLASS (Global Land Atmosphere System Studies) panel. The objective of this study is the comparison of models that simulate water, energy, and CO_2 cycles with continuous observations at five different sites. The proposed experiment has unique characteristics. PILPS San Pedro-Sevilleta not only focuses on a different environment than previous PILPS experiments, but it also will employ appropriate system methods for parameter estimation, that will help the modeling groups to identify parameter sets that make the models consistent with the data.

The availability of 4+ years of data at two locations with similar vegetation coverage but hundreds of kilometers apart provides an exciting opportunity for cross-validation of the model results and for comparison of different models. The three different vegetation types existing at the data sites also provide a quick look at the diversity of environments in arid lands and will make it possible to determine whether or not further distinction is required to better represent the water, energy, and CO_2 exchanges taking place over such areas.

In previous PILPS studies *[Lettenmaier et al., 1996; Nijssen et al., 2003],* it was shown that the calibration of model parameters yielded improvement in the models performance. For this reason, we propose to use the multi-criteria framework and a set of optimization codes for calibration of hydro-meteorological models that has been developed and successfully applied to a variety of land surface models at the University of Arizona [*Gupta et al, 1998, 1999; Bastidas et al., 1999, 2001, 2002; Vrugt et al., 2003*]. This framework constrains the parameter estimation of land surface models to be consistent with observations and will allow for a comparison of "optimal" model performances. However, the use of this multi-criteria framework is not compulsory and the participants may carry out parameter estimation in the way they see fit.

Some of the science questions to be addressed by the PILPS San Pedro-Sevilleta experiment are:

- What is the ability of the models to reproduce the water, energy, and carbon exchanges in semi-arid environments?
- Are the current (usually single) representations of semi-arid lands in the models enough to reproduce the different environments that exist in those areas?
- Does model calibration reduce the among-model range in the model simulations?
- How much influence does the model parameterization have on the parameter estimations of "physically meaningful" parameters?
- Do current carbon representations, developed for forests, properly reproduce carbon exchanges over vegetated arid lands?

PILPS San Pedro-Sevilleta is open to models with and without a representation of carbon fluxes. To guarantee comparisons under similar conditions, all participants will be able to carry out calibrations/optimizations that do not use carbon flux information. Modeling groups that represent carbon processes will be required to perform an additional set of calibration and simulation experiments to evaluate the changes and potential *improvements* due to inclusion of the carbon information.

1.2 *Description of the sites and instrumentation*

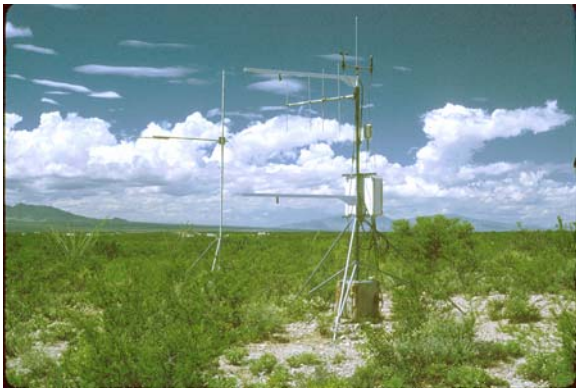
The proposed experiment will be carried out at five different sites located within the semi-arid Southwest USA, in the states of Arizona (3 sites) and New Mexico (2 sites) (See Figures 1 and 2 for locations). Two of the sites, Lucky Hills and Sevilleta Shrub, have a shrubby vegetation coverage with predominant species Acacia (*Acacia constricta*), tarbush (*Flourensia Cernua*), creosote brush (*Larrea divaricata*), and desert zinnia (*zinnia pumila*). The Kendall and the other Sevilleta site are grasslands with predominant species sideoats grama (*Bouteloua curtipendula*), black grama (*Bouteloua eriopoda*), hairy grama (*Bouteloua hirsuta*) and lehmann lovergrass (*Eragrostis lehmanniana*). The Tucson site has shrubs, grass, and saguaro cacti.

The data for the Lucky Hills and Kendall sites has been collected by the USDA-ARS Tucson from January 1997 till December 2000 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. The data from the Tucson site was collected by Jim Shuttleworth's group of the University of Arizona from May 1993 to June 1995 using an eddy covariance system on a 9 m high tower *[Unland et al, 1996]*. The measurements are of sensible and latent heat, and soil temperature. The data at the Sevilleta sites was collected by Eric Small of the University of Colorado using towers with a height of 10 m. Measurements include sensible, latent, and CO_2 fluxes, soil temperature and soil moisture at 5 cm depths.

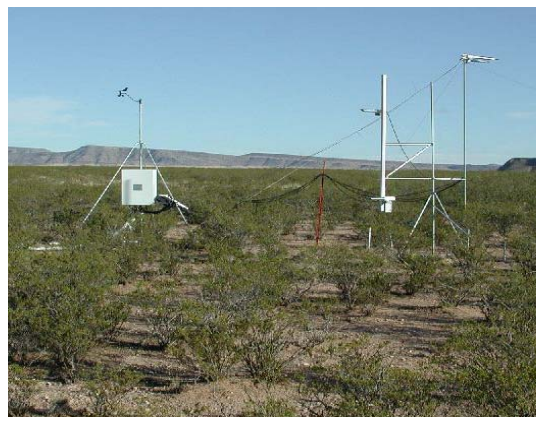
Table 1. PILPS San Pedro-Sevilleta experimental sites

Site	Longitude West	Latitude North	Elevation [m.a.s.l.]	Precipitation [mm/year]	Annual Temperature [°C]
Lucky Hills Shrubland	110°03'05''	31°44'37"	1372	340	18.6
Kendall Grassland	109°56'28"	31°44'10"	1526	340	19.3
Tucson Shrub/cacti	111°49'48"	32°13'01"	730	305	20.2
Sevilleta Grassland	106°43'30"	34°20'30"	1730	270	17.2
Sevilleta Shrubland	106°44'39"	34°20'05"	1776	270	16.9

Measurements of the vegetation coverage and descriptions of the soil types at all the sites are also available. The soils tend to be coarse loams with high clay content. The detailed information will be provided to the participants.




Lucky Hills and Kendall sites

Tucson and Sevilleta sites

Grass at Sevilleta site.

Figure 1. PILPS San Pedro-Sevilleta experimental sites

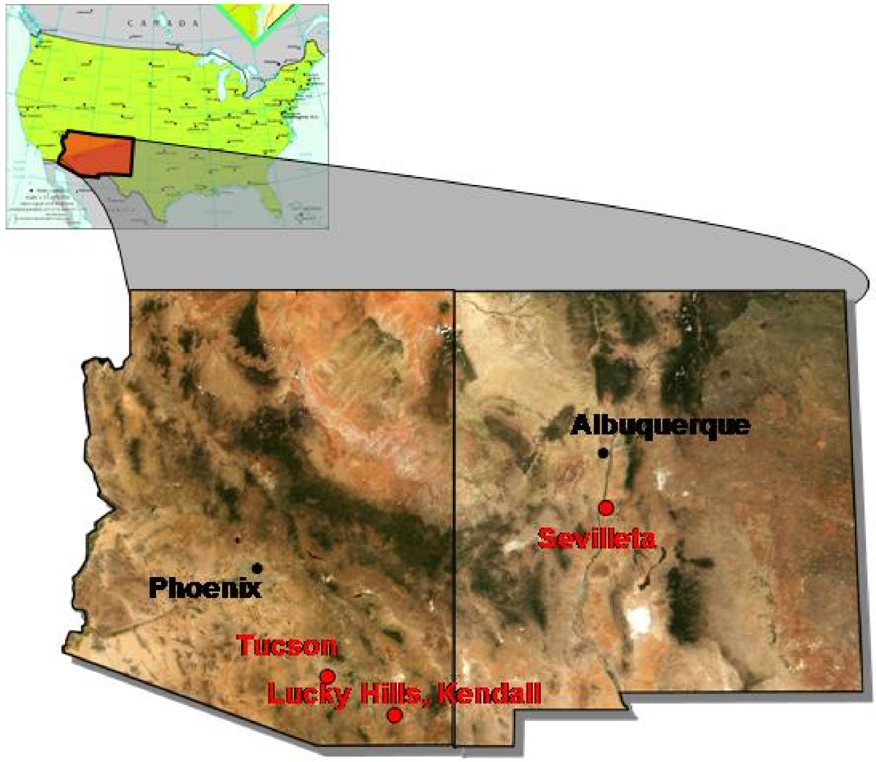


Figure 2 Location of Data Sites

2 EXPERIMENTAL SETUP

2.1 *General*

The experiment is focused on data obtained from observational towers that are located comparatively near to each other (1-5 km) to hundreds of kilometers apart, but within similar environments. We propose a set of offline experiments that will allow for a series of “cross-validations” or evaluations of model performance. As an innovation, we propose the use of optimization routines for the identification of “calibrated/optimal” parameter sets for all the models. The model intercomparison will be among the “optimal” performances of the models. The optimizations will be carried out within the multi-criteria framework developed at the University of Arizona, which will provide the computer codes and training for this exercise. The participants are not obliged to use this framework for their parameter estimation procedures, if they so choose. To assess the impact of calibration in the model performance a “*default*” parameter set and the associated model run will be requested. The evaluation of the models will be carried out at annual, monthly, daily, and 30 minute time scales.

For the initialization of the models 6 months of forcings will be provided at the Kendall and Lucky Hills sites. Observations of soil moisture at 4 different depths but uneven time intervals exist for the year 1996 at both locations. The data will be provided as ancillary information for initialization. The data were collected at the same locations. At the Sevilleta sites forcing will be provided from the NLDAS. For the Tucson site, 15 months of data will be provided and the participants will be requested to use the initial 3 months as initialization data. It is believed that due to the dry conditions existing at the data locations the data supplied for initialization will suffice.

2.2 *Proposed intercomparison runs*

All participating model groups will be requested to complete the following sets of model experiments. Models that do not simulate carbon fluxes and stores will only complete the experiments in Set A. Each of these experiments is explained in more detail in the following sections

Set A Non-carbon simulations

1. Default model parameters
   The default parameters (pre-calibration) will be based on the default model parameters for semi-arid regions, combined with a description of general conditions at the sites. No information regarding moisture and energy fluxes will be provided at this stage. These simulations will form the baseline simulations with which the results from the calibrated simulations will be compared.
2. Ad-hoc calibrated model parameters
   Specified model parameters will be calibrated by each model group using the calibration methods that each modeling group normally employs. For some groups this means manual calibration, while others may employ automated calibration procedures.
3. Multi-criteria calibrated model parameters
   Specified model parameters will be calibrated by each model group using the multi-criteria calibration framework developed at the University of Arizona. This last step is optional, but we strongly encourage participants to participate. The results from this step will act as a check on how well the ad-hoc calibration procedures perform, and will allow us to more directly compare “optimal” model performance.

Set B Carbon simulations

Models that include the representation of carbon stores will repeat the same series of model simulations as those in set A, but with the carbon component enabled. Those models in which the carbon component cannot be disabled, will only perform the simulations in set B. Their results will still be compared with the other models in set A, but in the resulting PILPS publications it will be emphasized that these models simulate carbon by default in addition to energy and moisture fluxes.

The Lucky Hills and Kendal sites will be used for temporal “split sample” tests of model performance. Both sites have data available for a 4 year period, including a “wet” and a “dry” year (1998 and 1999). At the same time, the availability of the New Mexico Sevilleta sites, with similar soil and vegetation characteristics, but hundreds of kilometers away, allow for spatial split sample testing and for evaluation of parameter transferability.

2.2.1 *Single location temporal "split-sample" test.*

The participants will be provided with all the forcing data and a subset of the evaluation data from the Lucky Hills and the Kendall sites. The evaluation data sets will contain wet and dry periods for the calibration of their models. Each modeling group will be requested to run the model for the full 4 year period for each site, using the default and calibrated model parameters.

The models will be compared using the provided data subset, the non-provided subset, and the whole set at each of the locations. The Tucson site data set will also be provided to the participants to check their parameter estimates with different vegetation coverage. The participants will be provided with the forcings for all the periods. However, the outputs to be used for evaluation of the model calibration, i.e. latent and sensible heat fluxes, CO_2 fluxes, ground temperatures, and soil moisture, will be provided only for the non-evaluation (calibration) periods. This part of the experiment will help in establishing the consistency of the parameter estimation procedures and the consistency of the models under different forcing conditions.

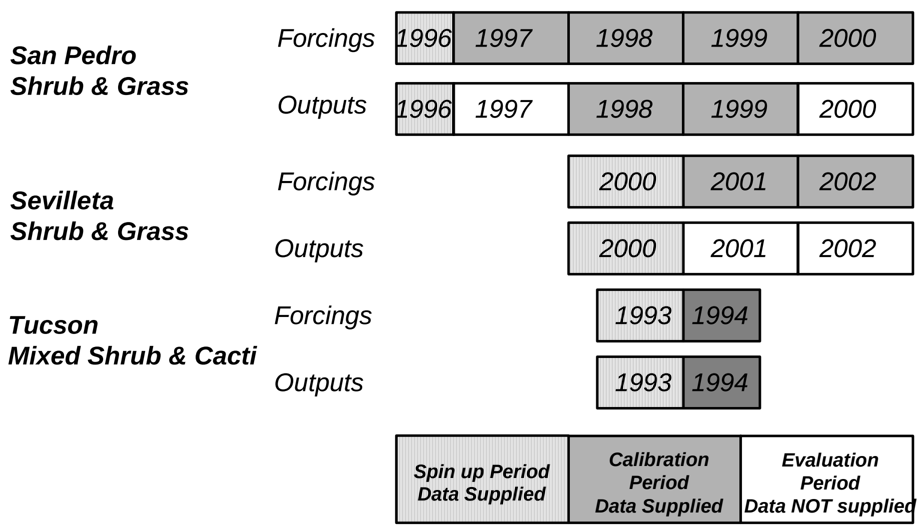


Figure 3. Calibration and evaluation periods

2.2.2 *Spatial "split-sample" test, transferability of parameters*

The Sevilleta shrub and grassland sites will be used to evaluate and compare the model performances based on the parameters obtained from the Lucky Hills and Kendall sites respectively. This exercise will check for the assumed general behavior of arid lands in the models and for the transferability of parameter estimates in similar but spatially distant conditions. The availability of soil moisture measurements at the

Sevilleta sites will allow for the testing of the consistency in the model estimation of state variables that were not used for the parameter estimation procedures.

2.2.3 *Carbon flux simulations*

A separate evaluation will be carried out for the carbon simulations using the same schemes of "split-sample" tests, i.e. temporal and spatial. Only models that simulate carbon fluxes will be requested to perform this runs. These tests will only be carried out at the Lucky Hills and Kendall sites.

3. MODEL FORCINGS

All the data will be provided using the NetCDF format and the ALMA conventions (www.lmd.jusssieu.fr/ALMA).

3.1 *Surface forcings*

The surface forcings will be provided with a 30 minute time step for all the sites, except Tucson, where the time step is 20 minutes. They include:

- Rainfall and snowfall.
- Wind speed.
- Air temperature.
- Specific humidity, derived from relative humidity at the Lucky Hills and Kendall sites.
- Incident shortwave radiation
- Incident longwave radiation, from N-LDAS.
- Surface pressure form NCEP model outputs.

The corresponding variable names in ALMA conventions are: Rainf, Snowf, Wind, Tair, Qair, SWdown, LWdown, PSurf respectively. The values will represent backward averages.

3.2 *Ancillary data*

The texture characteristics of the soil at each of the sites will be provided. In addition, values of a number of "observable" variables will be provided (as text files):

- Vegetation type
- Vegetation cover fraction
- Height of vegetation
- Leaf Area Index
- Surface albedo
- Longwave emissivity

4. MODEL OUTPUT

4.1 Output variables

Table 2 shows the list of ALMA variables that each model group should return. Note that this list may be adjusted in the final set of instructions that will be distributed through the PILPS San Pedro-Sevilleta web site (www.sahra.arizona.edu/pilpssanpedro). Variables that are not produced by a model should simply be omitted from the returned data files. Flux variables should be provided as backward-averages over the model timestep, while state variables should be provided as instantaneous values at the end of the model timestep. See the ALMA convention for details.

Table 2. Variables to be returned (see ALMA web site for definitions, units, and details)

ALMA Variable name	Description
O.1 General energy balance components	
SWnet	Net shortwave radiation
LWnet	Net longwave radiation
Qle	Latent heat flux
Qh	Sensible heat flux
Qg	Ground heat flux
Qa[1]	Advective energy
DelSurfHeat	Change in heat storage
DelColdCont[1]	Change in snow cold content
O.2 General water balance	
Snowf[1]	Snowfall rate
Rainf	Rainfall rate
Evap	Total evapotranspiration
Qs	Surface runoff
Qsb	Subsurface runoff
DelSoilMoist[1]	Change in soil moisture storage
DelSWE[1]	Change in snow water equivalent
DelSurfStor	Change in surface water storage
DelIntercept	Change in interception storage
O.3 Surface state variables	
SnowT[1]	Snow surface temperature
VegT	Vegetation canopy temperature
BaresoilT	Temperature of bare soil
AvgSurfT	Average surface temperature
RadT	Surface radiative temperature
Albedo	Surface albedo
SWE[1,2]	Snow water equivalent
SurfStor	Surface water storage
O.4 Subsurface state variables	
SoilMoist[3]	Average layer soil moisture

SoilTemp[3]	Average layer soil temperature
SoilWet	Total soil wetness
O.5 Evaporation components	
ECanop	Interception evaporation
TVeg	Vegetation transpiration
ESoil	Bare soil evaporation
EWater	Open water evaporation
RootMoist	Root zone soil moisture
CanopInt	Total canopy water storage
ACond	Aerodynamic conductance
O.6 Other hydrologic variables	
WaterTableD	Water table depth
O.8 Variables to be compared with remote sensing data	
LWup	Upward longwave broadband radiation
O.9 Carbon budget	
GPP	Gross primary production
NPP	Net primary production
NEE	Net ecosystem exchange
AutoResp	Autotrophic respiration
HeteroResp	Heterotrophic respiration
TotSoilCarb	Total soil carbon
TotLivBiom	Total living biomass

[1] Although very little snow falls at the PILPS San Pedro-Sevilleta sites, modeling groups are requested to return the snow variables to allow the computation of energy and water balances.
[2] Only total grid cell SWE needs to be reported
[3] 3D variables

4.2 Additional information

In addition to the model results, each group will be requested to return the parameter sets that were used for the simulations, as well as a description of the model and a list of references for the model.

4.3 File naming conventions

The file naming convention will be similar to that used in previous PILPS experiments. All model results for a single run at a single site will be returned as a single file with the name [modelname]_[simulation]_[location]_pilpssanpedro.nc. The [modelname] will be a model identifier unique to each model. The [simulation] will indicate which of the simulations specified in Section 2 is contained in the file, and can take the values "a1", "a2", "a3", "b1", "b2", "b3". Finally, the [location] refers to the code for each of the locations as follows:

- Lucky Hills shrubland = "lucky",
- Kendall grassland = "kendall",
- Tucson shrub/cacti = "tucson",
- Sevilleta grassland = "sev_grass",

- Sevilleta shrubland = "sev_shrub".

Thus, for example, the file "zzz_a1_kendall_pilpssanpedro.nc" will include all non-carbon base simulation made with the "zzz" model for the Kendall grassland site.

The parameters associated with each model simulations should be returned as a simple text file, with one file per location and model run. The naming convention for these files will be [modelname]_[simulation]_[location]_pilpssanpedro.par, where [modelname], [simulation], and [location] as defined above. For example, "zzz_a1_kendall_pilpssanpedro.par" will contain the model parameters used to produce the model results in "zzz_a1_kendall_pilpssanpedro.nc".

4.4 Expected data volumes

A modeling group that would complete all simulations would return no more than 30 data files and 30 parameter files. Because these will be point simulations, individual file sizes will be limited to a few MB, and all results for an individual model will fit on single CD-ROM. Specific instructions for returning model results will be provided on the PILPS San Pedro-Sevilleta web site.

The output information should include the results for all the five sites using a "default" parameter set and the estimated parameter sets. The corresponding parameter sets will also be requested.

5. PROPOSED ANALYSIS

As stated above the proposed experiment will attempt to test the models under the so called "split sample" framework and to establish the possible advantages of using ad-hoc and formalized parameter estimation procedures. The evaluation will include comparisons of the model outputs to the observations at the same site, but for a different time period; and comparisons at different locations, with similar physical characteristics.

The analysis will be carried out for the fluxes and state variables directly measured at the sites, and namely: latent heat flux, sensible heat flux, net shortwave radiation, net longwave radiation, soil temperature and soil moisture @ 5 cm depth, and the carbon flux.

Each of these variables will be compared to the observations for the following conditions using measures as the correlation coefficient, the Nash-Sutcliffe Efficiency, the root mean square error, the bias, the maximum distance, etc. In particular we will focus on the:

- Monthly mean
- Daily mean
- Daily amplitude
- Daily phase
- Min and max of the diurnal cycle
- Values at each time step

Based on the different error measures, we will attempt to estimate the usefulness of the parameter estimation procedures for the models. Optimization codes, in F90 and C, will be provided to the participants and training will be provided as part of a PILPS San Pedro-Sevilleta workshop to be held in Tucson, Arizona in August 27-29, 2003. For this evaluation, the performances of the models using the *default* parameter sets will be used as benchmarks.

We will work within the framework proposed by Klemes (1986) for model evaluation, i.e. the split sample test will be used for both temporal and spatial evaluations. In addition to that, variables not used in the calibration procedures will be used for performance evaluation.

6. DATA PROTOCOLS

All data handling and format requirements will follow the ALMA-3 guidelines, as described in the ALMA website (www.lmd.jussieu.fr/polcher/ALMA/dataex_main.html). Model results that do not conform to this convention will not be accepted.

6.1 *Data distribution and return*

The meteorological forcing data and the outputs will be provided via FTP, WWW, or CD as the participants choose. The output variables that a given model cannot provide or does not produce should simply be omitted in the netCDF file. The specific instructions about the sites and the naming conventions for the files will be provided via a web page that will be established and maintained at the University of Arizona.

6.2 *Quality control*

ALMA has made a screening program available to check the correctness of the output netCDF files prior to return and to ensure that the models conserve water and energy. This program based on those defined for the PILPS 2e Experiment will apply the annual water, energy, and carbon balance criteria; as well as ensuring that all variables are within reasonable ranges. The range requirements are not meant to comment on the appropriateness of model output, merely to verify unit correctness and sign. A number of utilities are freely available for plotting netCDF files, as listed on the ALMA web site, and we encourage their use as well. The screening program will be run after submission prior to any analysis. Any data that fail the screening will not be considered. Consequently, we encourage participants to run the program prior to submitting results.

6.3 *Results documentation*

Besides the model runs results some additional information will be requested from the participants:

- Short description of the model and the model structure. Include references.
- Description of the calibration procedure and which were the outputs used in the calibration.

- General impressions and comments on the results obtained based on the experience with the participants own models.
- Default parameter set (without calibration). Which parameters were calibrated and the calibrated parameter sets.
- Specific problems or concerns experienced.
- Details of any modification to the provided information.

7. EXPERIMENT STEPS

STEP 1 SAN PEDRO BENCHMARK RUN – NO CALIBRATION

Period: July 14, 1996 - December 31, 2000
Data from: Lucky Hills and Kendall
Data provided:

- Forcings, as stated in paragraph 3.1 of the experiment description
- Daily Average temperature for the 96 years of record at Tombstone, Arizona
- Soil moisture at different depths for year 1996. Time intervals not regular. They are provided to enable the model initialization.

Report back

- Period January 1st, 1997 – December 31st, 2000
- *"Default"* parameter values for semi-arid areas (without calibration)
- Backward averages at hourly time steps of all the variables listed in table 2
- Any special assumption

STEP 2 SAN PEDRO AD-HOC CALIBRATION

Period: July 14, 1996 - December 31, 2000
Data from: Lucky Hills and Kendall
Data provided:

- Data from Step 1
- Average daily albedo (10:00 – 15:00 hours)
- Tucson site observations
- NDVI weekly
- Greenness fraction

Report back

- Period January 1st, 1997 – December 31st , 2000
- Manually calibrated parameter values for semi-arid areas.
    - Only one set valid for both locations.
    - A specific set for each location
- Backward averages at hourly time steps of all the variables listed in table 2
- Any special assumption

STEP 3 SAN PEDRO AUTOMATIC CALIBRATION – SINGLE LOCATION TEMPORAL SPLIT SAMPLE TEST

Period: July 14, 1996 - December 31, 2000
Data from: Lucky Hills and Kendall
Data provided:

- Data from Step 2
- Observations for January 1st, 1998 – December 31st, 1999
  - Sensible Heat
  - Latent Heat
  - Ground Heat
  - Soil Temperature
  - CO_2 flux

Optimization Procedure

- Algorithm: MOSCEM, to be provided in Fortran 90 and C.
- Error function: Root Mean Square Error (RMSE) using 1 hour averages.
  - RMSE and Bias computations will be automatically implemented in the optimization code to be provided
- Optimize on:
  - Sensible Heat Flux, Ground Heat Flux
  - Sensible Heat Flux, Ground Heat Flux, Soil Temperature
  - Sensible Heat Flux, Ground Heat Flux, Soil Temperature, Carbon Flux
- Parameter bounds to be provided where possible
  - Optimize only time invariant parameters
    - Soil parameters
    - Vegetation parameters
    - Do not optimize geometric parameters
    - Do not optimize boundary conditions
- Pick best solutions based on
  - Automatically done by the optimization code provided
  - L_2 norm for RMSE
  - L_2 norm for Bias
  - Zero bias

Report back

- Period January 1st, 1997 – December 31st, 2000 for both locations (Kendall and Lucky Hills)
- Parameter sets
  - Entire Pareto set of 250 parameter sets
  - Best solutions
    - L_2 norm for RMSE
    - L_2 norm for Bias
    - Zero Bias
- Backward averages at hourly time steps of all the variables listed in table 2 for the
  - Entire Pareto set of 250 parameter sets

  - Best solutions
    - L_2 norm for RMSE
    - L_2 norm for Bias
    - Zero Bias
- Any special assumption

STEP 4 SEVILLETA BENCHMARK RUN – NO CALIBRATION

Period: January, 2001 - December, 2002
Data from: Sevilleta Grass and Sevilleta Shrub
Data provided:

- Forcings, as stated in paragraph 3.1 of the experiment description
- Daily Average temperature for the years of record at Albuquerque, New Mexico
- Parameters to be used: "*default*" (uncalibrated)

Report back

- Period January 2001 – December 2002 for both locations (Kendall and Lucky Hills)
- Backward averages at hourly time steps of all the variables listed in table 2 using "*default*" parameter values for semi-arid areas (before calibration)
- Any special assumption

STEP 5 PARAMETER TRANSFER FROM LUCKY HILLS (SHRUB) AND KENDALL (GRASS) TO SEVILLETA SHRUB AND SEVILLETA GRASS SITES. SPATIAL SPLIT SAMPLE TEST

Period: January 2001 - December 2002
Data from: Sevilleta Grass and Sevilleta Shrub
Data provided:

- Meteorological Data
  - Data from Step 4
- Parameters to be used:
  - Lucky Hills from step 3 in Sevilleta Shrub
  - Lucky Hills from step 3 in Sevilleta Grass
  - Kendall from step 3 in Sevilleta Shrub
  - Kendal from step 3 in Sevilleta Grass

Report back

- Period January 2001 – December 2002 for both locations (Grass and Shrub)
- Backward averages at hourly time steps of all the variables listed in table 2 for the
  - Entire Pareto set of 250 parameter sets, i.e. 1000 solutions
  - Best solutions, i.e. 12 solutions
    - L_2 norm for RMSE
    - L_2 norm for Bias
    - Zero Bias solution

- Any special assumption

STEP 6 SEVILLETA AUTOMATIC CALIBRATION. SINGLE LOCATION TEMPORAL SPLIT SAMPLE TEST

Period: January 2001 - December 2002
Data from: Sevilleta Grass and Sevilleta Shrub
Data provided:

- Data from Step 4
- Observations for January 1st, 2002 – December 31st, 2002
    - Sensible Heat
    - Latent Heat
    - Ground Heat
    - Soil Temperature
    - Soil Moisture

Optimization Procedure

- Algorithm: MOSCEM
- Error function: Root Mean Square Error (RMSE) using 1 hour averages.
    - RMSE will be automatically implemented in the optimization code to be provided
- Optimize on:
    - Sensible Heat Flux, Ground Heat Flux
    - Sensible Heat Flux, Ground Heat Flux, Soil Temperature
    - Sensible Heat Flux, Soil Temperature, Soil Moisture
- Parameter bounds to be provided where possible
    - Optimize only time invariant parameters
        - Soil parameters
        - Vegetation parameters
        - Do not optimize geometric parameters
        - Do not optimize boundary conditions
- Pick best solutions based on
    - Automatically done by the optimization code provided
    - L_2 norm for RMSE
    - L_2 norm for Bias
    - Zero Bias

Report back

- Period January 1st, 2001 – December 31st, 2002 for both locations
- Parameter sets
    - Entire Pareto set of 250 parameter sets
    - Best solutions
        - L_2 norm for RMSE
        - L_2 norm for Bias
        - Zero bias

- Backward averages at hourly time steps of all the variables listed in table 2 for the
    - Entire Pareto set of 250 parameter sets
    - Best solutions
        - L_2 norm for RMSE
        - L_2 norm for Bias
- Any special assumption

STEP 7 ADDITIONAL QUESTIONS THAT MAY ARISE

To be determined. There is a strong possibility of application of the procedures for additional testing on Australian locations.

8. PROPOSED TIMELINE

- February 2004, Submission of final experimental protocol
- March 15th,2004, Submission of model information and default parameter sets used by the models for semi-arid areas
- March 15th, 2004, Distribution of forcing data to the participants through the website www.sahra.arizona.edu/pilpssanpedro
- March 31, 2004, Submission of default parameter simulations STEP 1
- April 1, 2004, Distribution of calibration data for STEP 2 and STEP 3
- April 30, 2004, Submission of results from STEP 2 – ad hoc calibrations
- April 30, 2004, Submission of results from STEP 3 – automatic calibrations
- June 1st, 2004, Distribution of data for STEP 4
- June 15, 2004, Submission of results from STEP 4
- June 30th, 2004, Submission of results from STEP 5
- July 1st, 2004, Distribution of data for STEP 6
- July 30th, 2004, Submission of results from STEP 6
- October, 2004, Workshop for analysis of preliminary results, at Utah State University.
- January, 2005, AMS session for final analysis and presentation of results

9. REFERENCES

Bastidas, L. A., H.V. Gupta, S. Sorooshian, W.J. Shuttleworth and Z.L. Yang "Sensitivity Analysis of a Land Surface Scheme using Multi-Criteria Methods", *Journal of Geophysical Research*, Vol. 104, No D16, p. 19,481-19,490, 1999.

Bastidas, L.A., H.V. Gupta, and S. Sorooshian, 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, American Geophysical Union, 2001.

Bastidas, L. A., H. V. Gupta, and S. Sorooshian, 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, 2002

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

Gupta, H.V., L. A. Bastidas, L., S. Sorooshian, W.J. Shuttleworth and Z.L. Yang, Parameter Estimation of a Land Surface Scheme using Multi-Criteria Methods", *Journal of Geophysical Research*, Vol. 104, No D16, p. 19,491-19,504, 1999.

Gupta, H.V., S. Sorooshian, and P.O. Yapo, Toward improved calibration of hydrologic models: Multiple and non-commensurable measures of information, *Water Resources Research*, *34* (4), 751-763, 1998.

Henderson-Sellers A., Pitman A., Love P., Irannjead P., Chen T. (1995) The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS): Phases 2 and 3. Bull. of the Amer. Met. Soc. 76:489-503

Henderson-Sellers A., Yang Z., Dickinson R. (1993) The Project for Intercomparison of Land Surface Parameterization Schemes. Bull. of the Amer. Met. Soc. 74:1335-1349

Klemes, V., 1986. Operational testing of hydrological simulation models, *Hydrological Science Journal*, 31(1-3), 13-24.

Pitman A., Henderson-Sellers A., Abramopoulos F., Avissar R., Garratt J., Frech M., Hahmann, A., Koster R., Kowalczyk E., Laval K., Lean J., Lee T., Lettenmaier D., Liang X., Mahfouf J.-F., Mahrt L., Milly P., Mitchell K., de Noblet N., Noilhan J., Pan H., Pielke R., Robock A., Rosenzweig G., Schlosser C., Scott R., Suarez M., Thompson S., Verseghy D., Wetzel P., Wood E., Xue Y., Yang Z.-L., , Zhang L. (1993) Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS). Results from off-line control simulations (Phase 1A). Gewex report IPGO Publication

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

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

Unland, H.E., P.R. Houser, W.J. Shuttleworth, and Z.L. Yang, 1996. Surface flux measurement and modeling at a semiarid Sonoran Desert site. *Agricultural and Forest Meteorology*, 82, 119-153.

Vrugt, J. A., H. V. Gupta, L.A. Bastidas, W. Bouten, and S.Sorooshian, A new effective and efficient algorithm for multi-objective optimization of hydrologic models, *Water Resources Research*, in press, 2003

Yapo, P.O., H.V. Gupta, and S. Sorooshian, Multi-objective global optimization for hydrologic models, *Journal of Hydrology*, *204*, 83-97, 1998.

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