Evaluation of Climate Model Predictions of Surface Variables from Hourly Observations at the ARM Southern Great Plains Site

Tom Phillips

September, 2007

Model outputs, observational data, and analysis methodology

This section includes background information on the forecast variables from the AM2 and CAM3 models that are analyzed in the present study, their relationship to observables at the ARM Southern Great Plains (SGP) site, and the analysis approach that is followed.

Procedures for climate model initialization and forecasting

Initial conditions for the AM2 and CAM3 model forecasts each day of the year 2000 were determined from 00Z ERA40 reanalyses of state variables mapped to each model's horizontal/vertical grid. Model forecasts of state variables then were output and archived for the subsequent 12Z-36Z period each day for the model atmospheric vertical column for that grid point lying closest to the ARM Central Facility (CF) at the Southern Great Plains site in Northern Oklahoma.

The ARM observations for the year 2000 were derived "continuous forcing" data. These observations were temporally interpolated so as to line up with the same 12Z-36Z period as the model forecasts, and the three-dimensional atmospheric variables also were interpolated from model levels to the same pressure levels as the corresponding ARM observations (personal communication, M. Fiorino and J. Hnilo). Thus, comparisons of model and observational variables *that are of the same type* can be made directly.

Available model outputs and ARM observational data

Because the AM2 and CAM3 are configured to output somewhat different variables, each model's predictions can only be directly compared with certain ARM observations. For example, the AM2 outputs screen-level (at a height of ~ 10 meters) eastward/northward wind components u/v at standard screen-level height (~10 meters) which can be directly compared with ARM observations; but the CAM3 only outputs winds at the lowest model ("lm") vertical level (at ~ 60 meters elevation). Also, while AM2 outputs both downward and upward components of the long-wave radiative fluxes, only the net surface long-wave flux is available for CAM3. On the other hand, CAM3 outputs supplemental information that is useful for the present analysis. Examples include the surface latent heat flux (in addition to the surface evaporative flux), and a vertically integrated total cloud cover fraction (in addition to cloud fractions at each model level). The similarities and differences between available ARM observations considered in this study and the respective model output variables are summarized in Table 1.

Table 1: ARM observations vs. AM2 and CAM3 output variables considered in this study, indicating the similarities and differences in their respective physical properties. An 'x' under the AM2 and CAM3 variable column denotes equivalency with the corresponding ARM observation, and 'lm' denotes a similar variable at the lowest model vertical level.

ARM Observations		AM2 variable	CAM3 variable
Notation	**Property (units)**		
cl	cloud area fraction in layer l (fraction)	x	x
clt	total (vertically integrated) cloud area fraction (fraction)		x
evs	surface evaporative flux (kg m^{-2} s^{-1})	x	x
hfls	surface latent heat flux (W m^{-2})		x
hfss	surface sensible heat flux (W m^{-2})	x	x
hurs	surface (2-meter) relative humidity (percent)	x	hur_lm
pr	total precipitation (kg m^{-2} s^{-1})	x	x
prw	precipitable water (kg m^{-2})	x	x
rlsd	surface downward long-wave flux (W m^{-2})	x	rls_net
rlsu	surface upward long-wave flux (W m^{-2})	x	
rssd	surface downward short-wave flux (W m^{-2})	x	x
rssu	surface upward short-wave flux (W m^{-2})	x	rss_net
tas	surface (2-meter) air temperature (deg K)	x	x
uas	surface (10-meter) eastward wind (m s^{-1})	x	ua_lm
vas	surface (10-meter) northward wind (m s^{-1})	x	va_lm
wspds	surface (10-meter) wind speed (m s^{-1})	x	

Analysis methodology

Our analysis in this study implicitly views the state variables and related fluxes at the land surface as *responses* to atmospheric *forcings*. For each model forecast, we first consider the downward short-wave and long-wave radiation as energy forcings, and then examine the response of the land surface in the form of the upward long-wave radiative and turbulent sensible heat fluxes, as well as the closely related surface air temperature. Attention next is directed toward the total (large-scale + convective) precipitation, as the chief hydrological forcing, and then to the related moisture response variables at the land-air interface such as the surface evaporative flux and humidity. Some mention also will be made of clouds which both modulate the atmospheric forcings and manifest part of the overall response, as well as precipitable water which represents the vertically integrated atmospheric-moisture response.

Admittedly, such an analysis approach over-simplifies the actual complexity of the coupled land-atmosphere system, wherein "forcings" cannot be cleanly distinguished from "responses". This approach also may invite a perspective that artificially separates the energy/moisture forcings and responses which in reality are different aspects of the same coupled system.

Particular attention thus will be directed to model predictions of surface evaporation, since this variable is a key "bridge" between the moisture and energy budgets. That is, land-surface evaporation is a function not only of the continental precipitation, soil moisture and runoff, but also of the surface net radiation and turbulent sensible heat transfer. As such, the prediction of surface evaporation is a telling indicator of multiple facets of model performance.

Model forecasts of surface variables during the year-2000 summer season

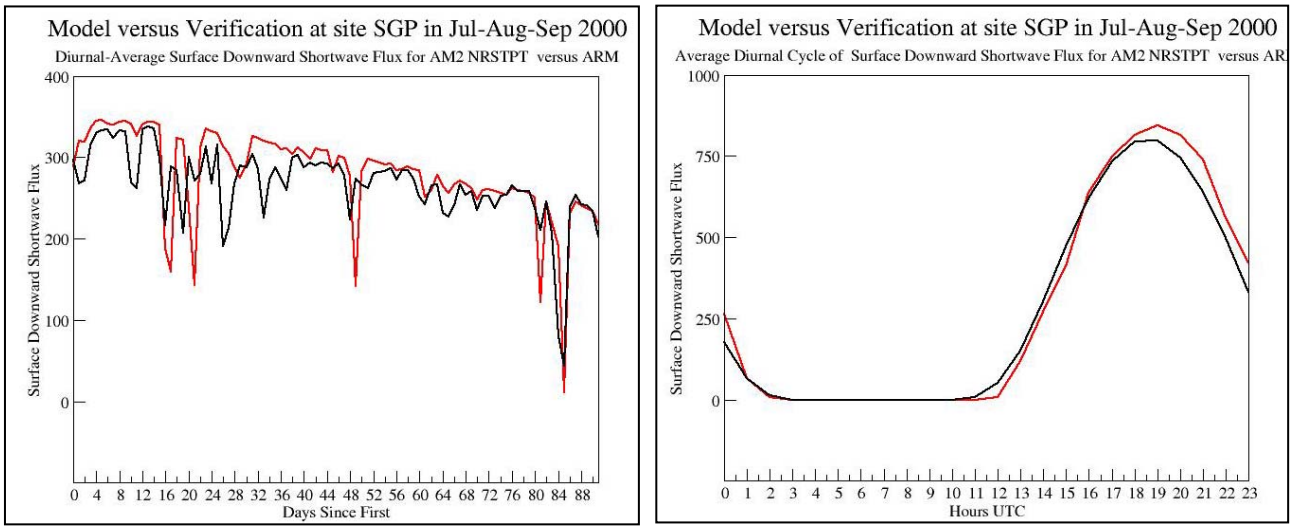
Because the model errors in surface variables as defined by ARM observations proved to be both the most stark and diverse during the year-2000 boreal summer, the greatest attention is directed to model forecasts during the July-August-September (JAS) summer season. A more cursory discussion of selected results in other seasons then follows in later sections.

AM2 forecasts during JAS

During JAS 2000, daily averages of downward short-wave radiative fluxes in AM2 show fairly good agreement with observations (Figure 1a), except for some notable lapses (e.g. overshoots/undershoots during July and the first part of August). The JAS-mean diurnal cycle of downward short-wave in AM2 (Figure 1b) also agrees well with observations, except that it peaks a bit later and displays a small positive bias during local afternoon hours.

Figure 1: Downward surface short-wave fluxes (units of W m^{-2}) during JAS of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are shown in a) as daily averages and in b) as the JAS-mean diurnal cycle. Note, 0 UTC corresponds to 6:00 p.m. local time, and 1200 UTC to 6:00 a.m. local time.

a) **b)**



These short-wave radiation errors are probably related to model shortcomings in representing the observed vertical profile of cloud cover, which seems to be generally under-predicted (Figure 2).

The AM2's reflected surface short-wave radiation (not shown) displays somewhat larger positive biases than does the downward short-wave flux (Figure 1). This presumably is because accurate prediction of the reflected surface short-wave requires not only a good representation of the downward short-wave radiation, but also that of the local surface albedo. The latter, in turn, depends both on a correct specification of local vegetation and, where bare ground is present, on a realistic prediction of soil moisture (i.e. wetter soil is associated with lower surface albedo).

Positive biases also are evident in AM2 daily average downward surface long-wave radiation (Figure 3a) which depends on the vertical profile of atmospheric moisture and clouds. An even more pronounced positive bias is evident in the daily average upward surface long-wave flux (Figures 3b) which is sensitive to the prediction of surface temperature and emissivity. Positive biases also are present in the respective JAS-mean diurnal cycles (not shown).

Figure 2: Vertical profiles of cloud amount fractions as a function of time (horizontal coordinate) during JAS of the year 2000, with ARM observations shown in the top panel, and AM2 forecasts below. Note that minute values of cloud fraction are present in the forecasts (indicated by violet shading in the bottom panel) even under essentially cloud-free conditions, due to numerical truncation errors associated with the AM2's finite-difference representation of atmospheric moisture.

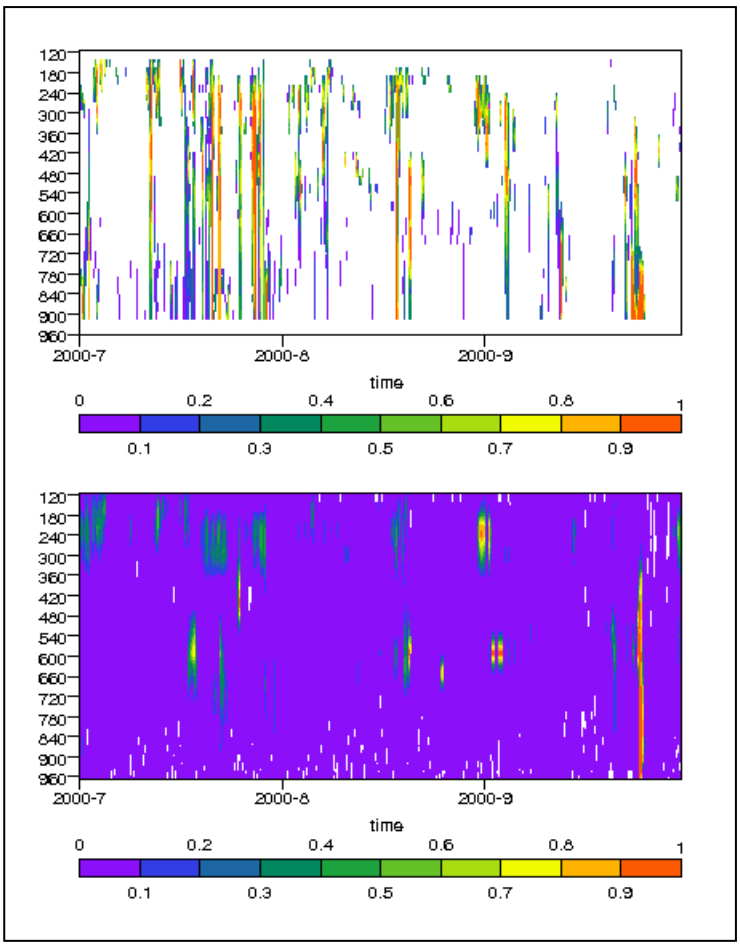
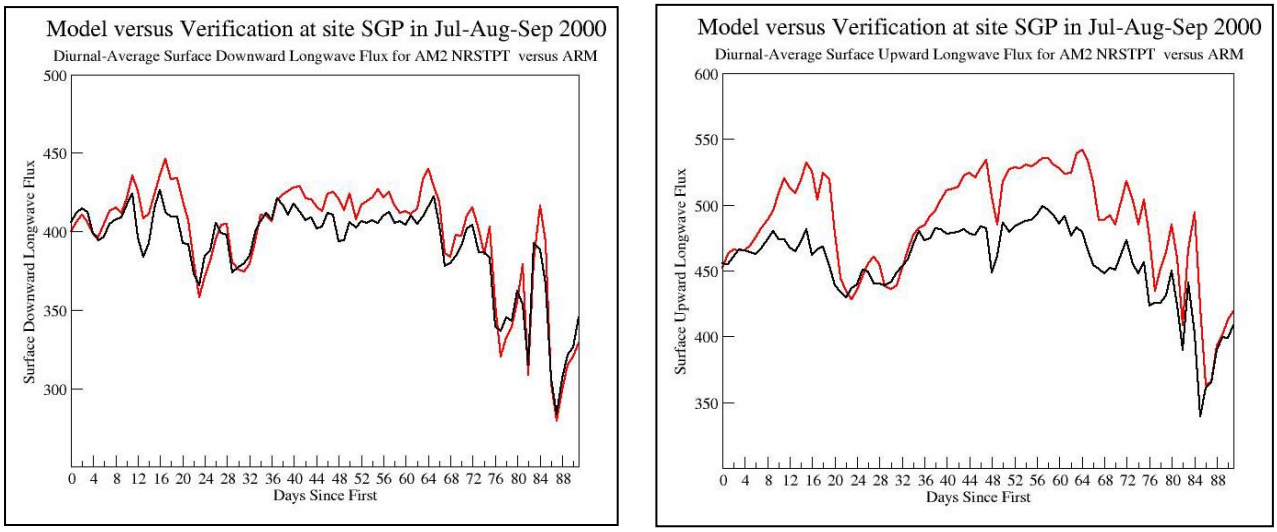


Figure 3: Time series of daily average surface long-wave fluxes (units of W m^{-2}) in JAS of the year 2000, directed a) downward, and b) upward for ARM observations (in black) vs. AM2 forecasts (in red).

a) **b)**

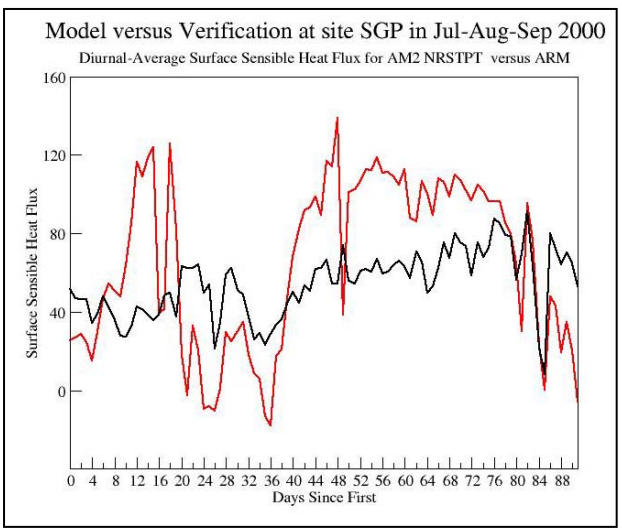


A fraction of the heating derived from the total net surface short-wave and long-wave radiative fluxes R_{net} is stored in the soil as a "ground flux" G, which usually is not directly observed, but can be estimated as a residual term. The difference $R_{net} - G$ then is available for partitioning into turbulent sensible and latent heat fluxes H and L at the surface, and the ratio H/L , denoted as the Bowen ratio, is a measure of that partitioning. Phase errors in the turbulent fluxes relative to observations thus often are indicative of an incorrect representation of the ground flux, and errors in the Bowen ratio may imply additional shortcomings in the land model (e.g. in its representation of evaporation from bare ground or from a vegetation canopy).

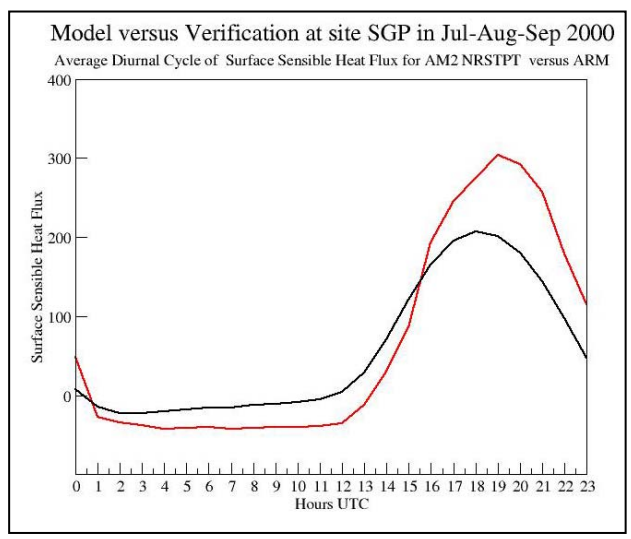
During JAS, AM2 surface sensible heat fluxes generally show positive biases in daily averages (Figure 4a) as well as in the daytime (but not nighttime) hours of the JAS-mean diurnal cycle (Figure 4b), where some phase error also is evident.

Figure 4: Surface (2-meter) sensible heat fluxes (units of W m^{-2}) during JAS of the year 2000 for observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle. In c) the scatter of hourly samples of model forecasts of sensible heat fluxes in AM2 vs. ARM observations is shown.

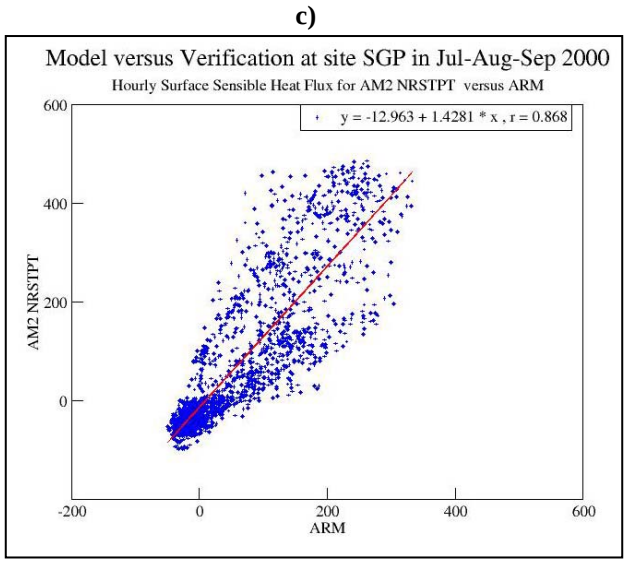
a)



b)



c)

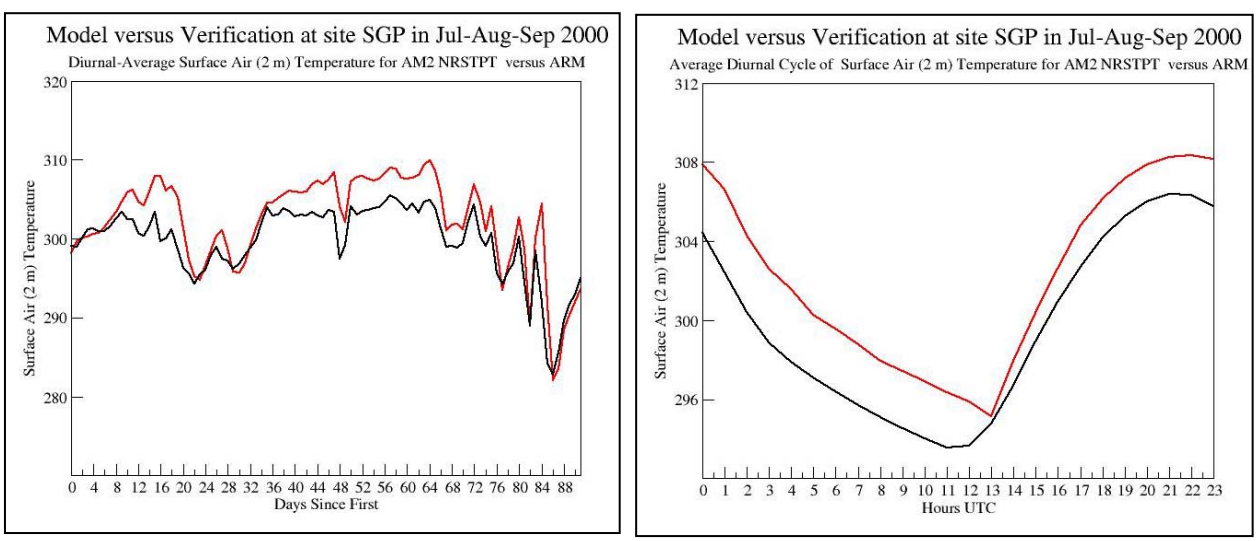


The bias and phase errors in the AM2 sensible heat flux are illustrated to a more fine-grained degree by a scatter plot of model forecasts vs. observations, where both datasets are sampled hourly (Figure 4c). Here the average bias error can be quantified in terms of the deviation of the regression line **y** = **mx + b** (in this case, **y** ~ 1.43**x –** 12.96) from the line **y = x** signifying perfect agreement of model forecasts and observations. In turn, a metric for the average phase error is the deviation of the correlation coefficient **r** of forecasts vs. observations (in this case, **r** ~ 0.87) from a value of **r =** 1.00, indicating perfect correlation. Note also that the higher the magnitude of the observed sensible heat flux in Figure 4c, the greater is the tendency of the forecast to scatter away from the regression line. This result suggests that, the larger the turbulent flux, the more problematical the use of a bulk aerodynamic scheme may be to approximate its magnitude.

The excessive AM2 upward long-wave flux is consistent with surface air temperatures that also are generally too high in both the daily averages (Figure 5a) and in the JAS-mean diurnal cycle (Figure 5b). The positively biased nighttime surface air temperature also is consistent with the negatively biased sensible heat flux of Figure 4b (i.e. a downward flux directed from an anomalously warm atmosphere to a cooler nocturnal land surface).

Figure 5: Surface (2-meter) air temperatures (units of degrees K) during JAS of the year 2000 for observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle.

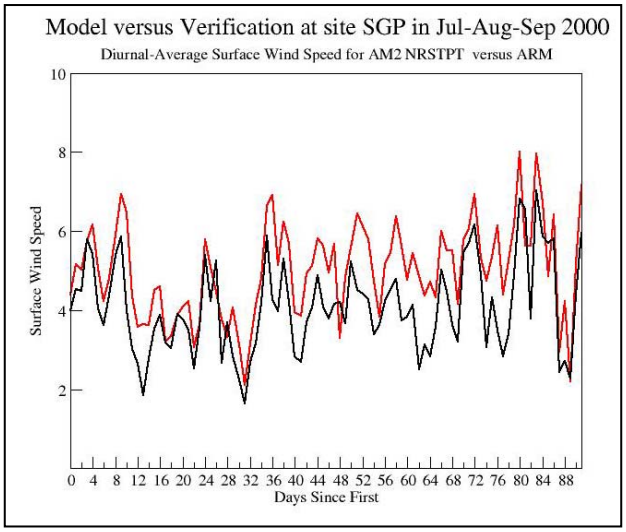
a) **b)**



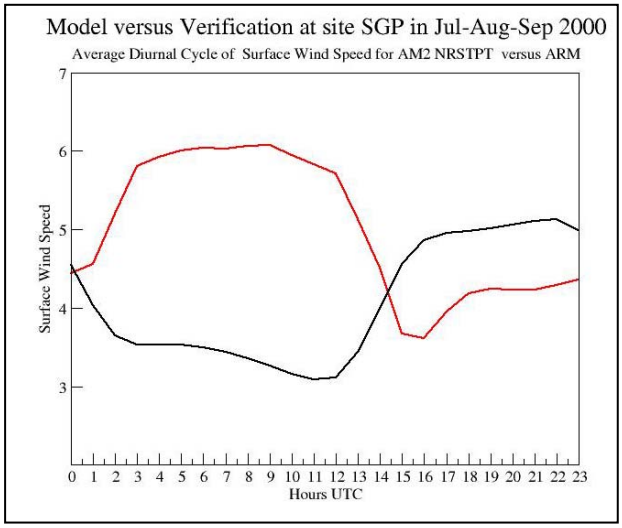
A positive bias is also present in daily averages of the AM2 surface dynamical state, e.g. in the surface wind speed (Figure 6a), which is a net result of overly strong nocturnal winds and overly weak daytime winds (Figure 6b). This anomalous diurnal cycle, which is mainly attributable to the v (northward-directed) component of the wind (Figure 6c), is typical of many atmospheric boundary layer (ABL) schemes used in today's GCMs (Svensson and Holtslag 2007).

Figure 6: Surface (10-meter) wind speeds (units of m s^{-1}) during JAS of the year 2000 for observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle. The JAS-mean diurnal cycle of the v (northward-directed) component of the surface wind is shown in c).

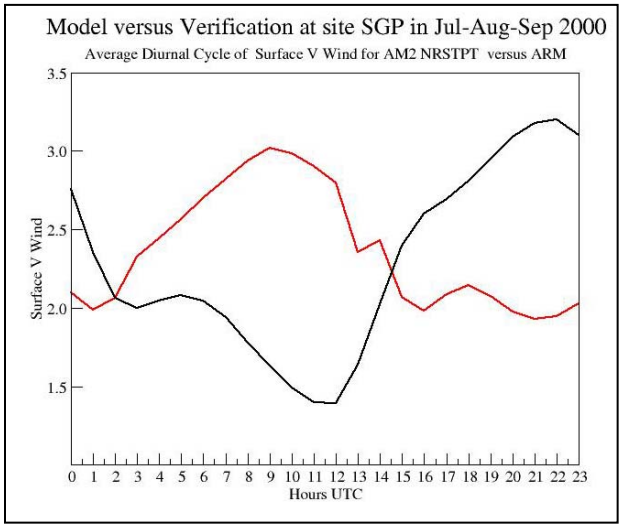
a)



b)



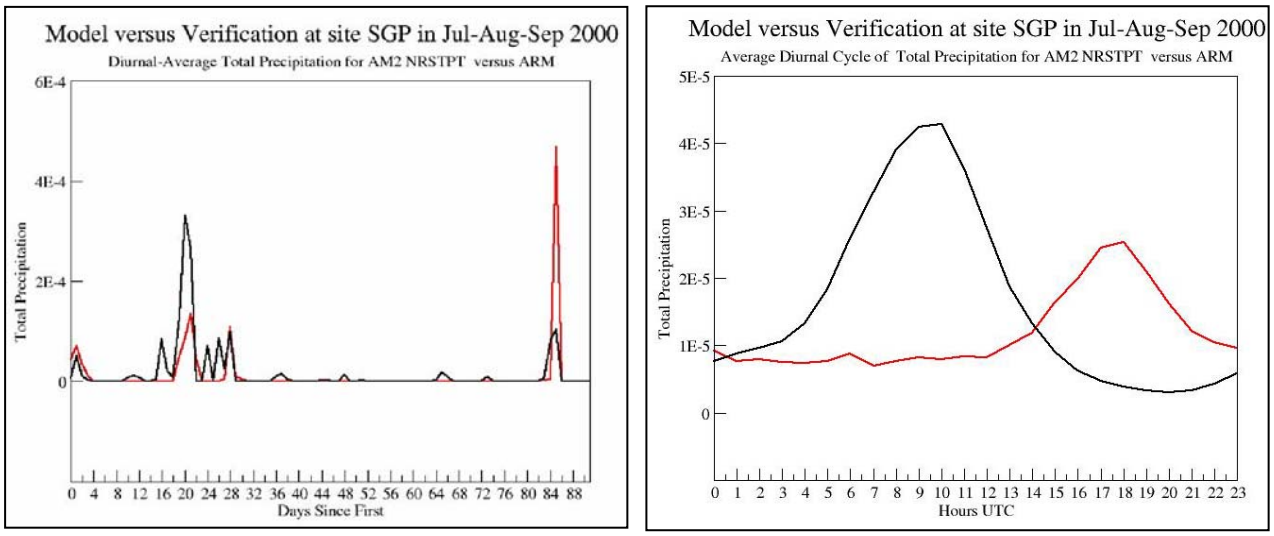
c)



Shifting attention next to moist processes, we consider total precipitation, the chief hydrological forcing for the surface. The AM2 forecasts the timing of daily precipitation events during JAS with some success (Figure 7a), but with general under-prediction of their magnitudes. Moreover, the model is unable to predict the observed diurnal cycle of JAS-mean precipitation (Figure 7b): instead of reaching a maximum a few hours after local midnight, as observed, AM2 precipitation peaks in mid-afternoon on average.

Figure 7: Total (large-scale + convective) precipitation fluxes (units of kg m^{-2} s^{-1}) during JAS of the year 2000 for observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle.

a) **b)**

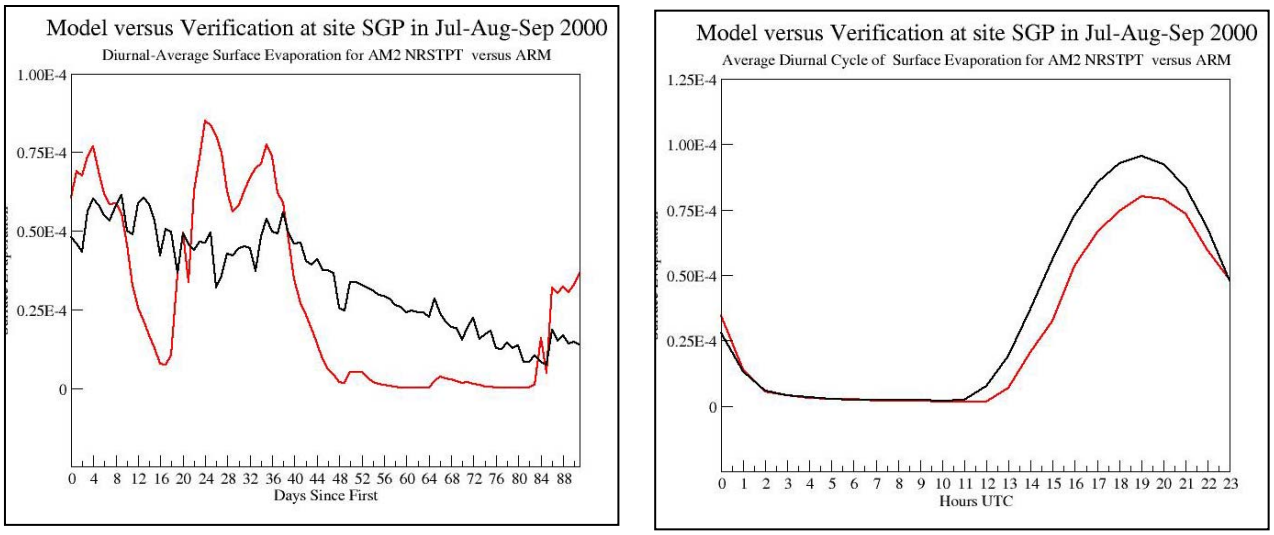


The magnitude of daily average surface evaporation also is generally under-predicted in AM2 forecasts (Figure 8a). However, between about days 22 to 40 the AM2 surface evaporation exceeds the observations, in spite of too-sparse predicted precipitation (and presumably, more desiccated soil) during this period. This behavior suggests that AM2 may overestimate land-surface evaporation upon recharge of soil moisture. On average, however, a negatively biased JAS-mean diurnal cycle of this surface flux occurs (Figure 8b), with some phase lag also evident (similar to that of the sensible heat flux in Figure 4b.

The scatter plot of hourly samples of observed and forecast surface evaporation (not shown) displays a pattern similar to that for the sensible heat flux in Figure 4c: model deviations from observations grow with the magnitude of the flux, suggesting that the tendency for the model to over-predict JAS surface evaporation immediately after precipitation events may have to do with the limitations of the bulk aerodynamic representation of surface fluxes as well as with shortcomings in the AM2's land model *per se*.

Figure 8: Surface (2-meter) evaporation fluxes (units of kg m^{-2} s^{-1}) during JAS of the year 2000 for observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle.

a) **b)**



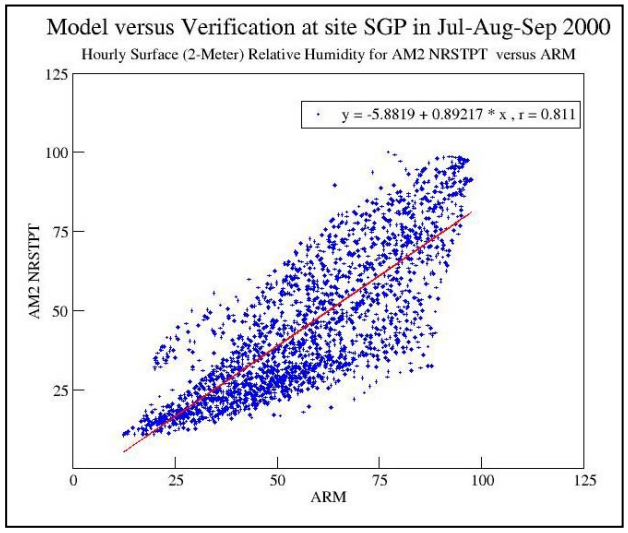
During JAS, precipitation over the U.S. Southern Great Plains mainly derives from convection. The ability to forecast summer precipitation is not only a test of the model's convective parameterizations, however, but also of its ability to correctly recycle the fraction of precipitation resulting directly from local surface evaporation. Thus, the sparse AM2 summer precipitation might be due as much to inadequate surface evaporation as to overly weak convection.

To investigate these physical mechanisms more closely, Klein et al. (2006) analyzed analogous AM2 forecasts near the ARM SGP site during June/July 1997, where model soil moisture boundary conditions were derived by imposing the observed local precipitation on the land model. Despite realistic magnitudes of surface evaporation, Klein et al. found that AM2 forecasts of precipitation were still too sparse relative to the observations. In view of the obvious additional erroneous intra-diurnal timing of convective precipitation, the AM2 shortcomings appear to be related to an inability to realistically propagate eastward-moving nocturnal convective systems in this region (Jiang et al. 2006).

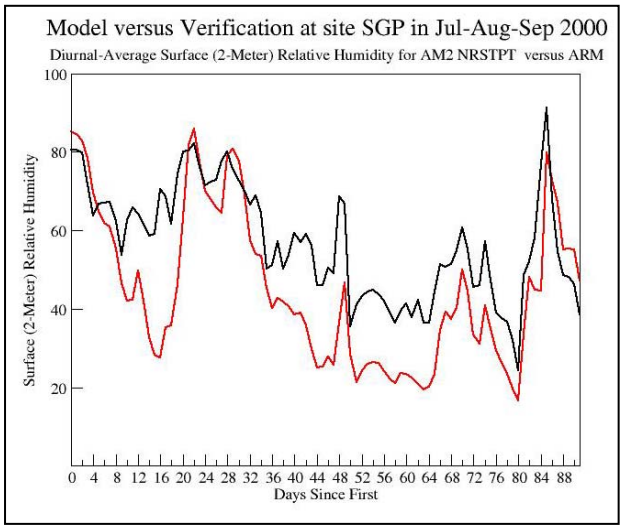
The scatter plot of hourly model samples vs. observations (Figure 9a) indicates a general negative bias (regression line $\mathbf{y} \sim 0.89\mathbf{x} - 5.88$), together with substantial phase error (correlation $\mathbf{r} \sim 0.81$). The scatter about the regression line also increases as relative humidity increases toward 100 percent, reminiscent of similar behavior in the scatter plot of surface evaporation. Negative bias errors also are evident in daily averages and JAS-mean diurnal cycle of surface relative humidity (Figure 9b,c). Moreover, the predicted diurnal cycle is too flat in amplitude, and it lags the observed phase, with larger deviations at night than during daytime.

Figure 9: Surface (2-meter) relative humidity percentages during JAS of the year 2000 are depicted in a) as a scatter of hourly samples of AM2 forecasts vs. ARM observations. Surface relative humidity percentages for observations (in black) vs. AM2 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the JAS-mean diurnal cycle.

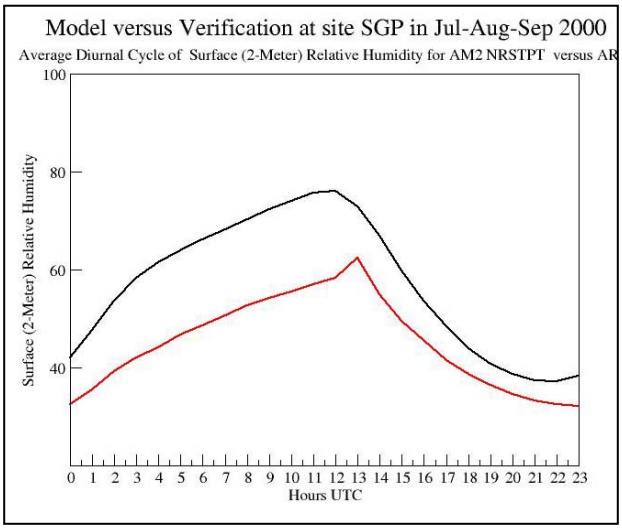
a)



b)



c)



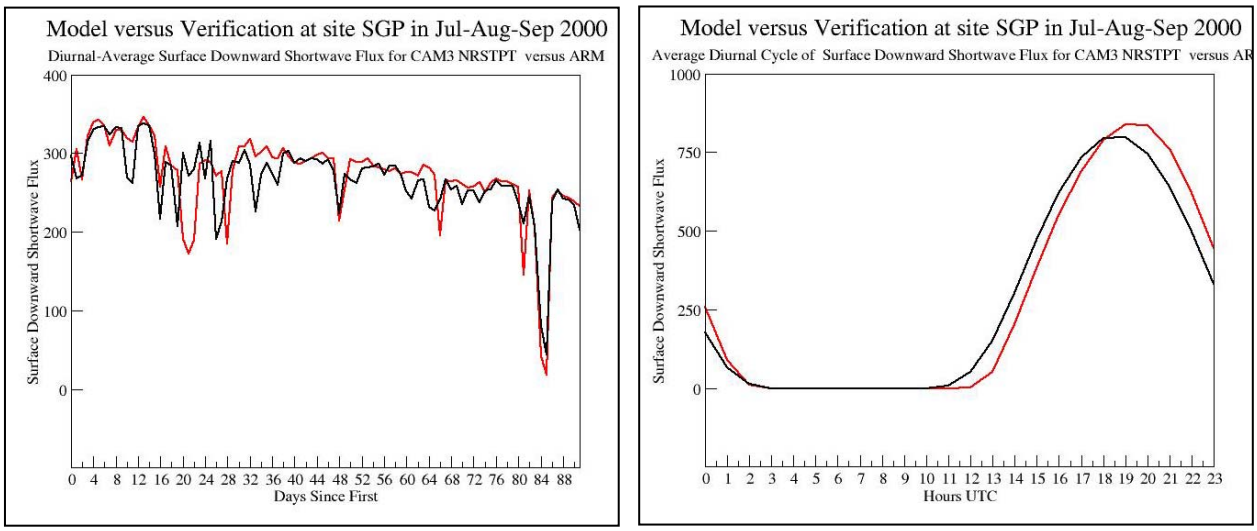
Despite the unrealistically low surface relative humidity, the AM2's daily average values of precipitable water (i.e. vertically integrated specific humidity--figure not shown) display little obvious bias relative to observations. However, the modeled JAS-mean diurnal cycle of precipitable water is unrealistic, in that its amplitude is too large, and its phase is also very different from that observed locally, peaking in mid-afternoon instead of early evening (figures not shown).

CAM3 forecasts during JAS

During the year-2000 JAS season CAM3 daily average downward short-wave radiation at the surface agrees fairly well with observations (Figure 10a), but the seasonal-average diurnal cycle (Figure 10b) shows some positive bias and a phase lag (Figure 10b).

Figure 10: Surface downward shortwave fluxes (units of W m^{-2}) during JAS of the year 2000 for observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle.

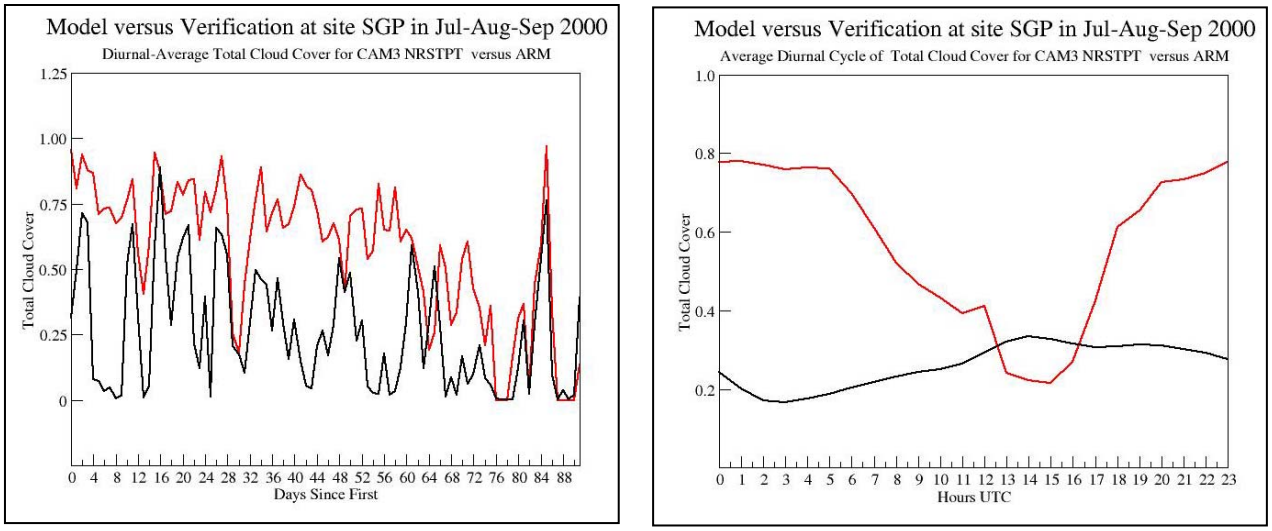
a) **b)**



This positive bias is surprising, given the CAM3's prediction of cloud cover almost every day which generally exceeds both the daily-average (Figure 11a) and the JAS-average intra-diurnal observations (Figure 11b); however, the *radiative* properties of the model clouds apparently are tuned to compensate, so that the downward surface short-wave flux ends up being about the correct magnitude. The JAS-mean diurnal cycle of the model's total cloud cover is also qualitatively different from that of the observations (Figure 11b), displaying much greater amplitude, and with a minimum value around 1400 UTC (~ 8:00 a.m. local time) when the observations show a relative maximum.

Figure 11: Total cloud cover fractions during JAS of the year 2000 for observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle.

a) **b)**

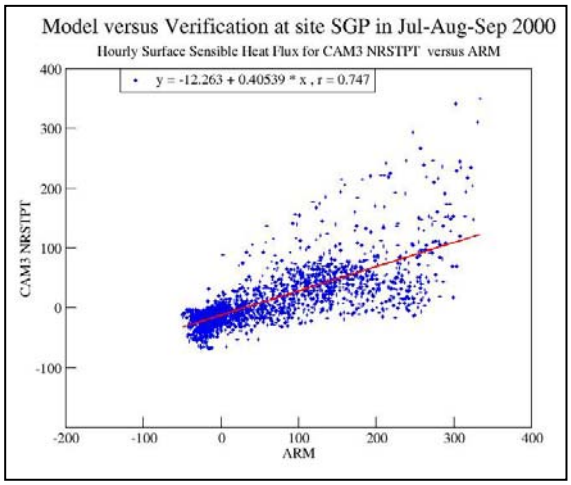


In the long-wave part of the radiation spectrum, only forecasts of net (upward minus downward) surface fluxes are available from CAM3. While daily averages of these fluxes are not obviously in error, the JAS-mean diurnal cycle of net long-wave is negatively biased during daylight hours, and is positively biased at night (figures not shown). A physical explanation (more fully elaborated below) is that anomalously low upward long-wave fluxes occur during daytime along with surface temperatures that are negatively biased due to an overly wet surface; but anomalously high nocturnal downward long-wave radiative fluxes result from too-large model cloud cover and humidity. Hourly samples also are biased negative (regression slope **m** ~0.61), with substantial phase error present as well (**r** ~ 0.69).

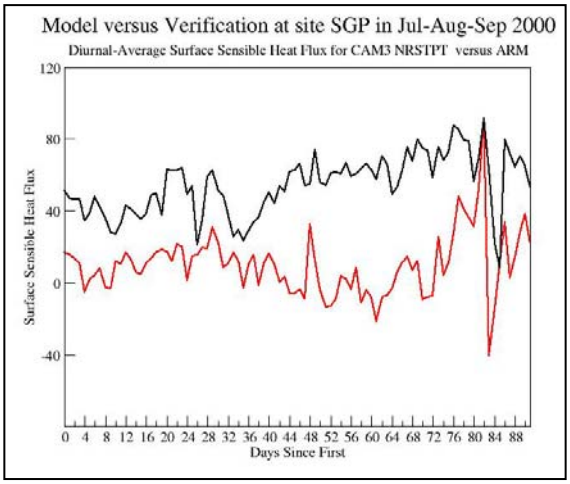
The CAM3 forecasts of surface sensible heat flux show large negative biases at all time scales (Figure 12), with some substantial phase errors as well (e.g. correlation **r** of hourly samples ~ 0.75).

Figure 12: Surface sensible heat fluxes (units of W m^{-2}) during JAS of the year 2000 are depicted in a) as a scatter of hourly samples of CAM3 forecasts vs. ARM observations. Surface sensible heat fluxes for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the JAS-mean diurnal cycle.

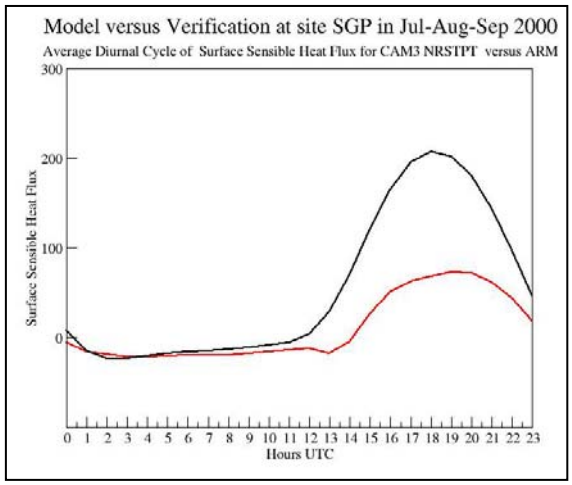
a)



b)



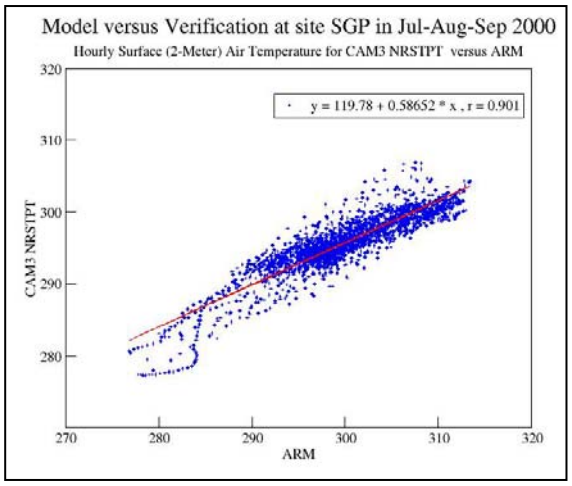
c)



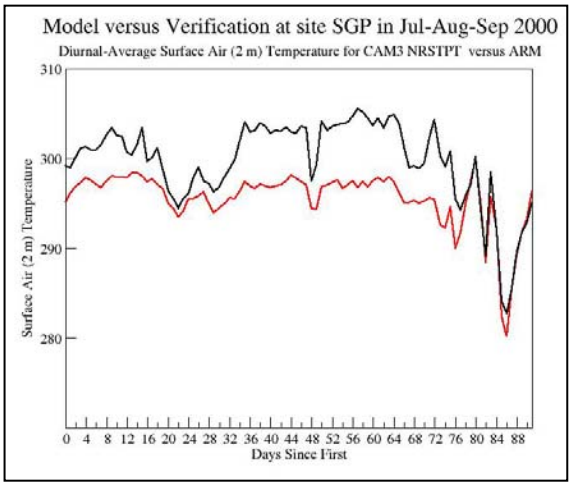
Surface air temperature also shows correspondingly negative biases at hourly (Figure 13a, regression slope **m** ~ 0.59) and daily average time scales (Figure 13b). From the JAS-mean diurnal cycle, the negative bias is seen to be more severe during daylight hours (Figure 13c).

Figure 13: Surface air temperatures (units of degrees K) during JAS of the year 2000 are depicted in a) as a scatter of hourly samples of CAM3 forecasts vs. ARM observations. Surface air temperatures for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the JAS-mean diurnal cycle.

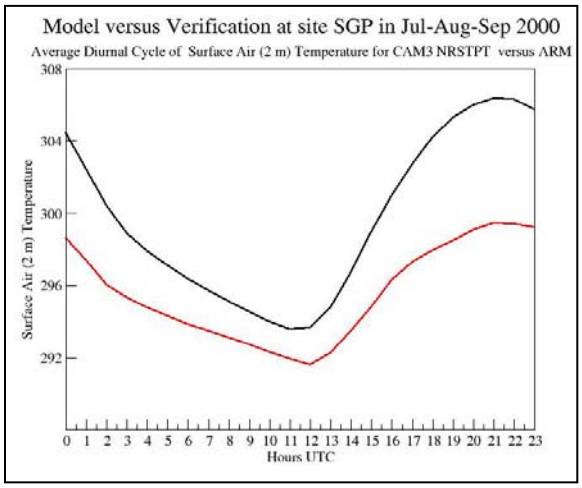
a)



b)



c)

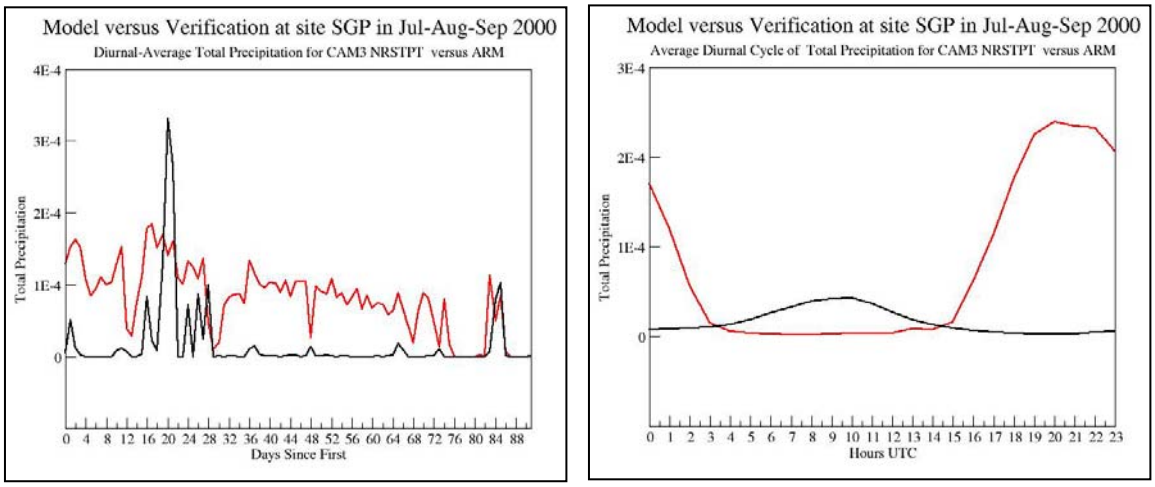


Although the CAM3's 10-meter surface winds are not available for direct comparison with ARM observations, the diurnal cycle of v-component winds at the lowest model level (~ 60 meter elevation) displays errors similar to those seen in AM2 (Figure 6).

In its forecasts of JAS hydrological variables, the CAM3 behaves very differently from the AM2 at the SGP site. Instead of producing too sparse precipitation, CAM3 rains almost every day (Figure 14a), in stark disagreement with observations. The CAM3'over-prediction of precipitation appears to be associated with a too frequently destabilized atmospheric column and subsequent triggering of the model's deep convection scheme (Williamson and Olson, 2007). Instead of predicting a few precipitation events as observed, CAM3 thus produces much "convective drizzle", a phenomenon that was also found in the previous CAM2 version of the model with similar convective parameterizations (Xie et al. 2004, Boyle et al. 2005, Williamson et al. 2005). Like AM2, precipitation in CAM3 also peaks in the afternoon, rather than at several hours past local midnight, as observed (Figure 14b).

Figure 14: Total precipitation fluxes (units of kg m^{-2} s^{-1}) during JAS of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle.

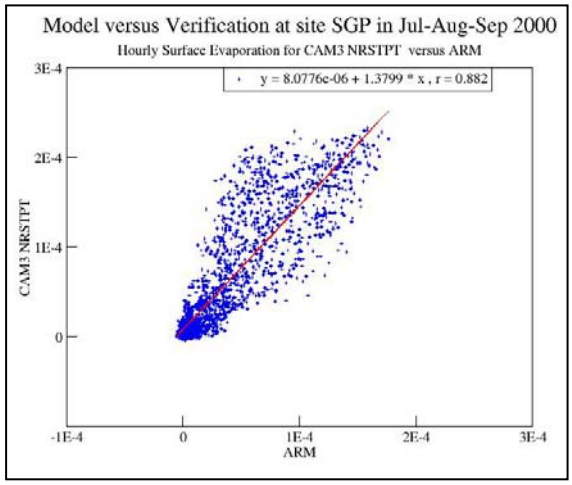
a) **b)**



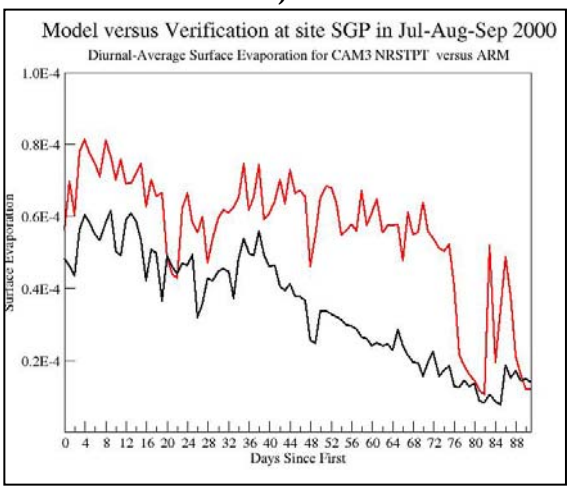
Consistent with the CAM3's excessive JAS precipitation, its surface evaporative flux is positively biased at all time scales (Figure 15).

Figure 15: Surface evaporation fluxes (units of kg m^{-2} s^{-1}) during JAS of the year 2000 are depicted in a) as a scatter plot of hourly samples of CAM3 forecasts vs. ARM observations. Surface evaporation fluxes for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the JAS-mean diurnal cycle.

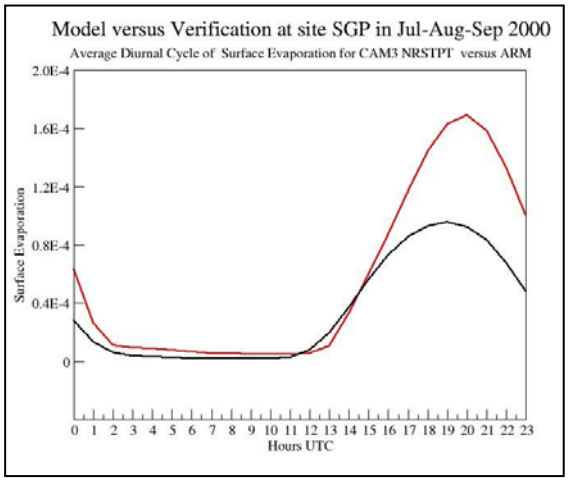
a)



b)



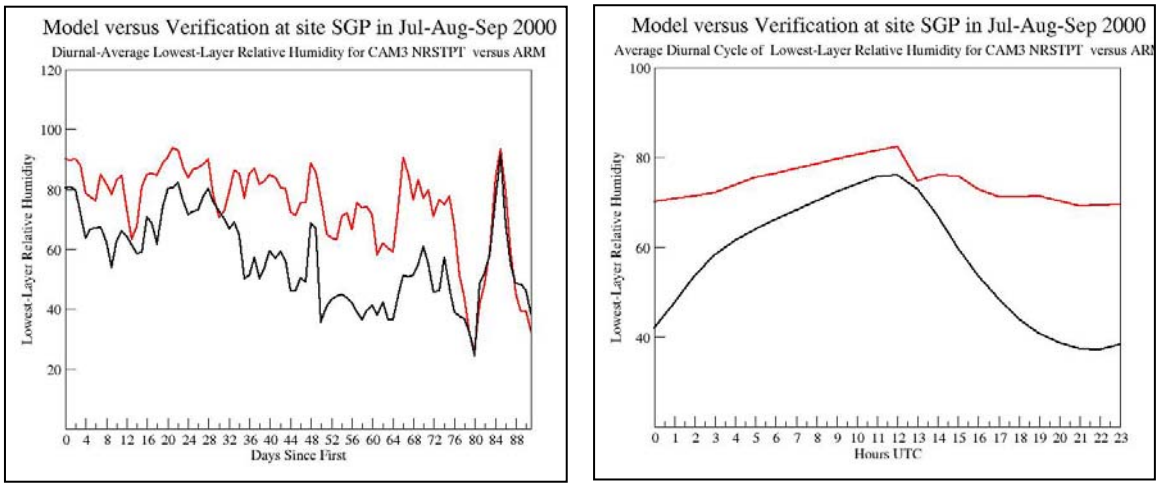
c)



Except toward the end of the season, the daily average relative humidity at the lowest vertical level of CAM3 (at ~ 60-meter elevation) also is positively biased relative to observations at 2-meter screen height (Figure 16a). (Not surprisingly, the CAM3 relative humidity at 60 m also shows much less intra-diurnal variation than the observations at 2 m in Figure 16b.) Despite the excessive surface humidity in CAM3, however, the daily average precipitable water is substantially less than observed (figure not shown), implying that the model atmosphere is too dry overall at the SGP site during JAS.

Figure 16: Lowest-level relative surface humidity percentage during JAS of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JAS-mean diurnal cycle.

a) **b)**



During JAS, the AM2 and CAM3 models therefore display qualitatively different behaviors in their forecasts of surface variables: AM2 shows a general propensity to run too dry, and CAM3 to run much too wet.

Model forecasts of surface variables during other seasons

The AM2 and CAM3 model errors during other seasons of the year 2000 are less egregious than in JAS, so our presentation of results will be more selective. We will focus on model biases that generally persist during entire seasons, and thus exemplify *systematic* deficiencies.

Forecasts during the year-2000 winter season

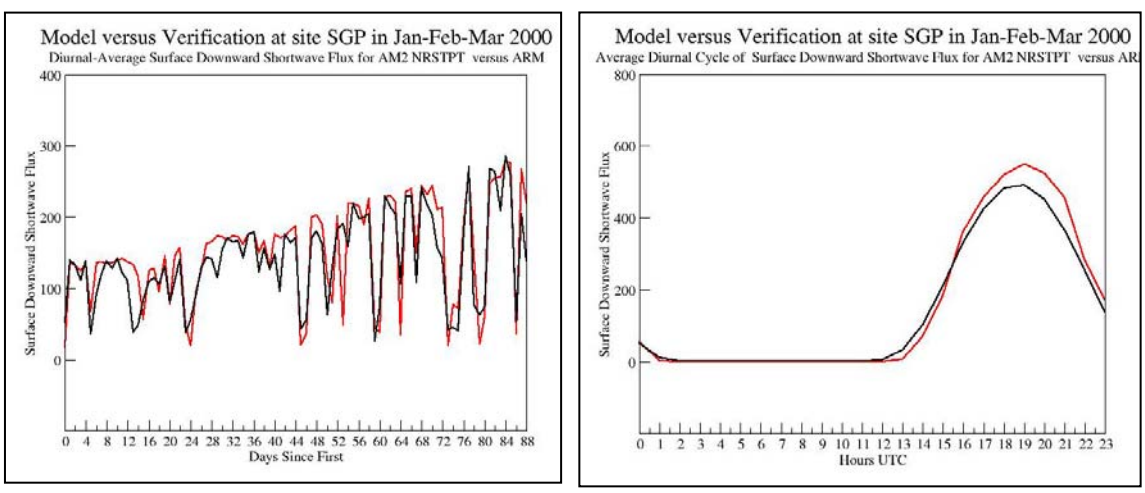
The AM2 and CAM3 forecasts during the year-2000 January-February-March (JFM) winter season afford the opportunity to evaluate model representations of surface variables for diametrically different synoptic conditions than those encountered in the JAS summer season. Plots illustrating the principal inter-seasonal differences are selectively presented below.

AM2 forecasts during JFM

As in JAS (Figure 1), the AM2 downward surface short-wave radiation is biased somewhat positive relative to observations (Figure 17a,b), while the upward surface short-wave flux (not shown) is biased negative, apparently because of too little reflected surface short-wave when snow covers the ground.

Figure 17: Surface downward short-wave radiation fluxes (units of W m^{-2}) during JFM of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

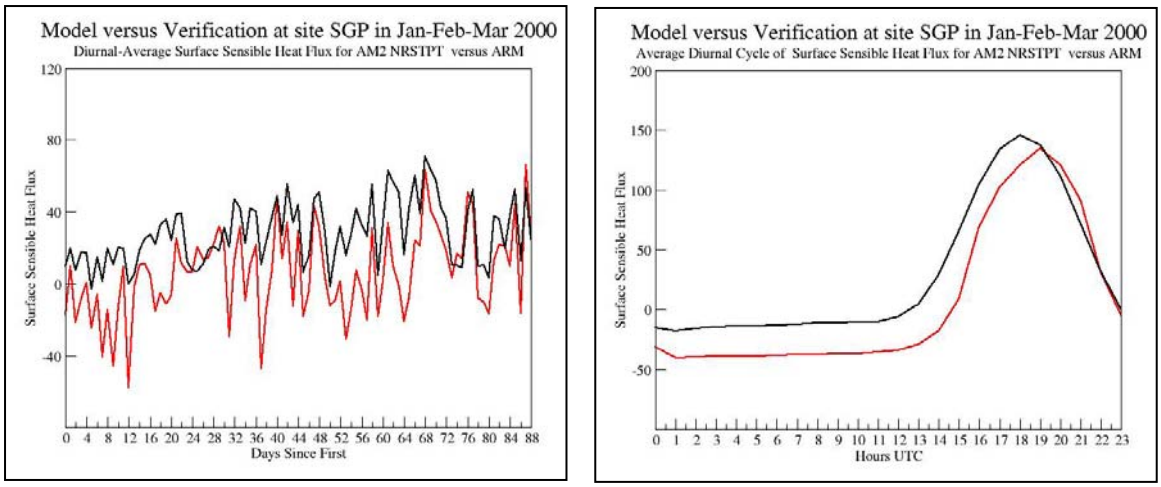
a) **b)**



The energy implications are that the AM2 land component absorbs somewhat more short-wave radiation than observed, and thus the model's ground storage of heat would tend to be unrealistically high. However, during daylight hours the downward long-wave flux is biased negative, thus offsetting this tendency. At night, the upward surface long-wave flux also is biased positive, making for greater radiative cooling of the land than observed (figures not shown). The resultant upward directed sensible heat flux is therefore negatively biased (Figure 18a,b).

Figure 18: Surface sensible heat fluxes (units of W m^{-2}) during JFM of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

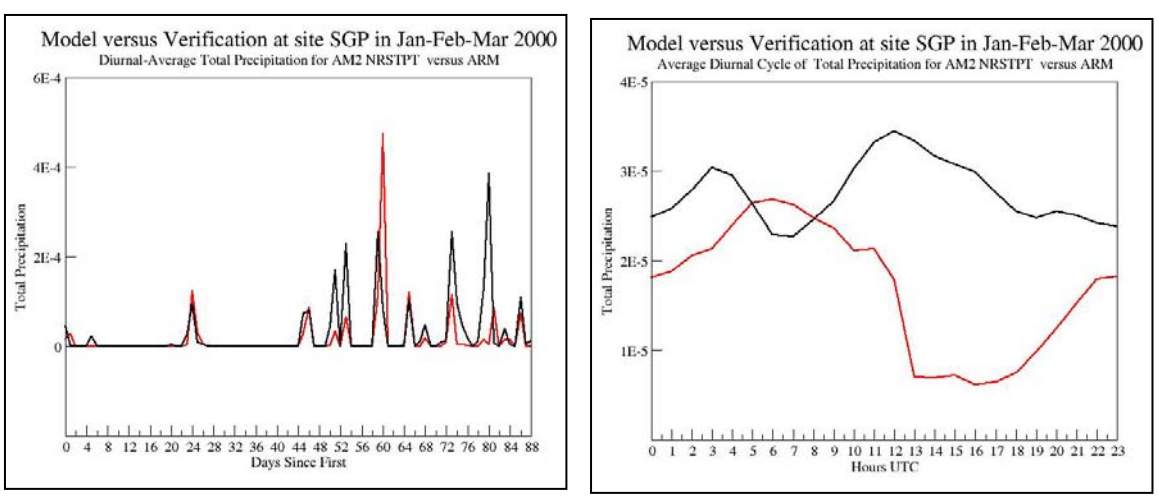
a) **b)**



During JFM, AM2 hydrological variables also deviate from observations. For example, while the timings of AM2 precipitation events are represented fairly well, their magnitudes are generally too low (Figure 19a). The model's seasonal-mean diurnal cycle of precipitation also is qualitatively different than that observed (Figure 19b).

Figure 19: Total precipitation fluxes (units of kg m^{-2} s^{-1}) during JFM of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red), are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

a) **b)**

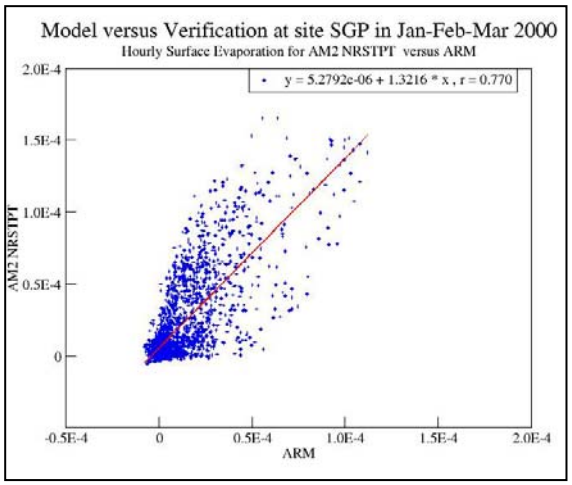


Despite the AM2's general under-prediction of precipitation magnitude, surface evaporation shows a large positive bias at all time scales (Figure 20). In JFM, the model thus predicts a Bowen ratio (sensible heating divided by latent heating) that is much too

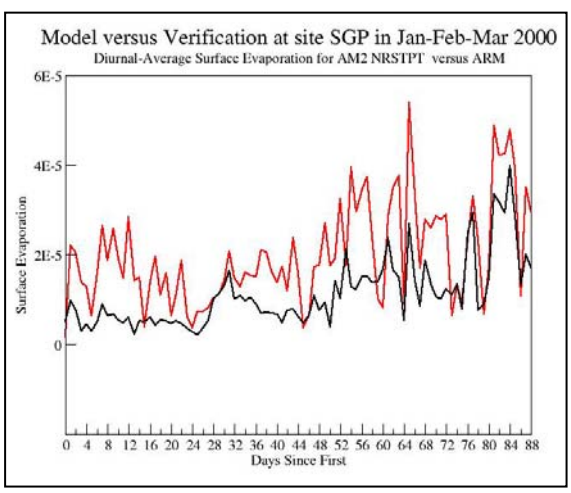
low. This discrepancy may be due either to inherent shortcomings in the land model or in the AM2 atmospheric bulk transfer schemes for estimating surface fluxes.

Figure 20: Surface evaporation fluxes (units of kg $m^{-2} s^{-1}$) during JFM of the year 2000 are depicted in a) as a scatter plot of hourly samples of AM2 forecasts vs. ARM observations. Surface evaporation fluxes for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the JFM-mean diurnal cycle.

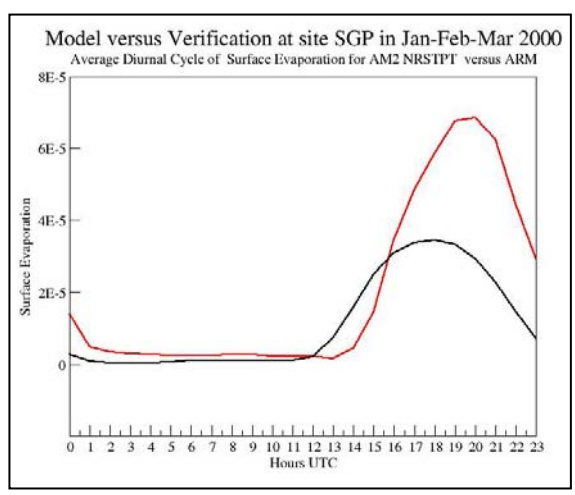
a)



b)



c)



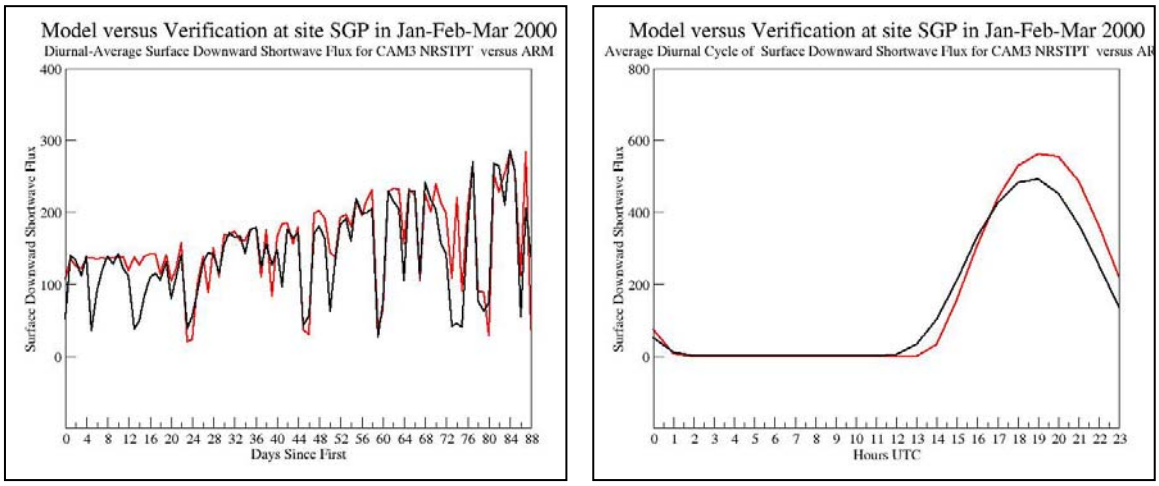
The model predictions of surface relative humidity also exhibit positive, but less extreme biases, while the amplitude of the JFM-mean diurnal cycle is lower than observed; however, the precipitable water displays the opposite behavior, indicating that overall the AM2 atmosphere is biased somewhat too moist during this season (figures not shown).

CAM3 forecasts during JFM

Like AM2, the downward surface short-wave flux in CAM3 shows a positive bias relative to observations (Figure 21). The net short-wave (downward minus upward) displays a similar bias as well (figure not shown).

Figure 21: Surface downward shortwave fluxes (units of W m^{-2}) during JFM of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

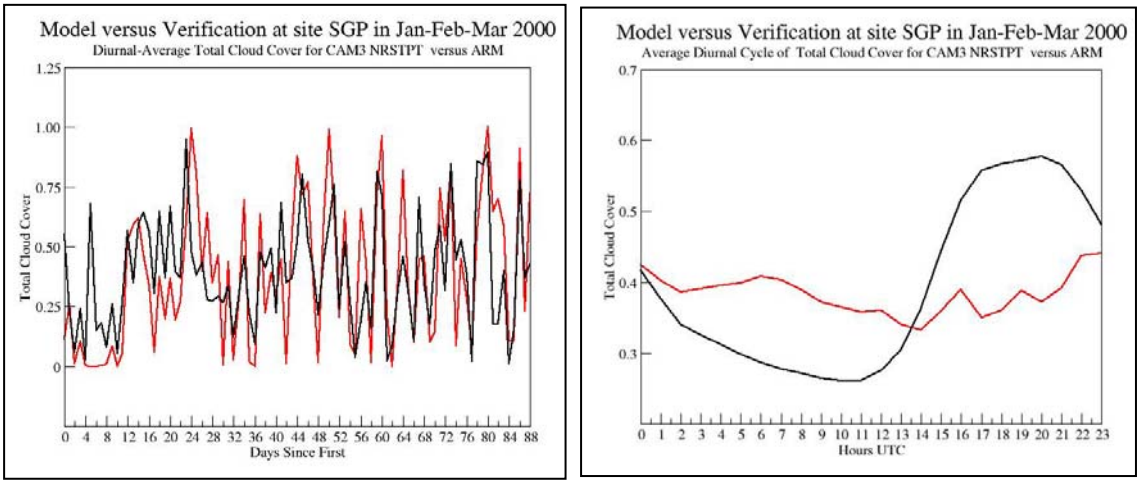
a) **b)**



Daily averages of the total JFM cloud cover in CAM3 are in fair agreement with observations, albeit with excessive inter-day variability (Figure 22a). This seeming agreement belies large errors in the amplitude and phase of the JFM-mean diurnal cycle (Figure 22b), where the CAM3's low total cloud cover during daylight hours is consistent with the positive bias seen in the surface downward short-wave radiation (Figure 21b).

Figure 22: Total cloud cover fractions during JFM of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

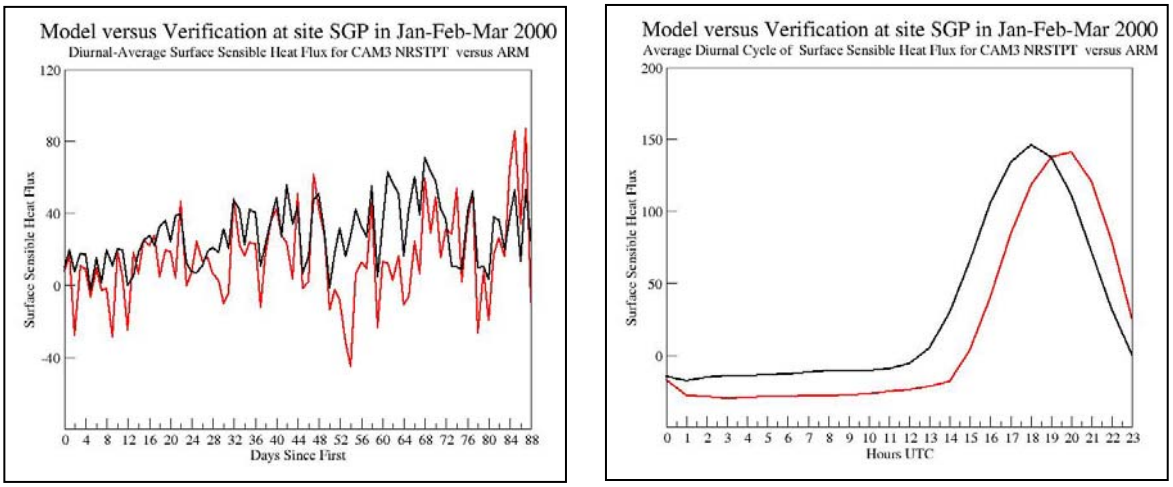
a) **b)**



However, the implied anomalous heating of the ground is offset by even larger biases in the surface net (upward minus downward) long-wave radiation (figures not shown), resulting in negatively biased ground temperatures. The sensible heat flux thus is depressed relative to observations at both hourly (regression slope **m** ~ 0.88) and at daily average time scales (Figure 23a). This negative bias is also present in the amplitude of the JFM-mean diurnal cycle (Figure 23b), especially at night when overly cold ground temperatures are predicted (not shown), producing a too-large downward-directed sensible heat flux. (Excessively cold wintertime near-surface temperatures are commonly found in global atmospheric models, e.g. Viterbo et al. 1999.) Moreover, the large phase lag seen in the diurnal cycle of sensible heat flux (Figure 23b) suggests that the actual ground heat storage is poorly represented by the CAM3 during JFM.

Figure 23: Surface sensible heat fluxes (units of W m^{-2}) during JFM of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

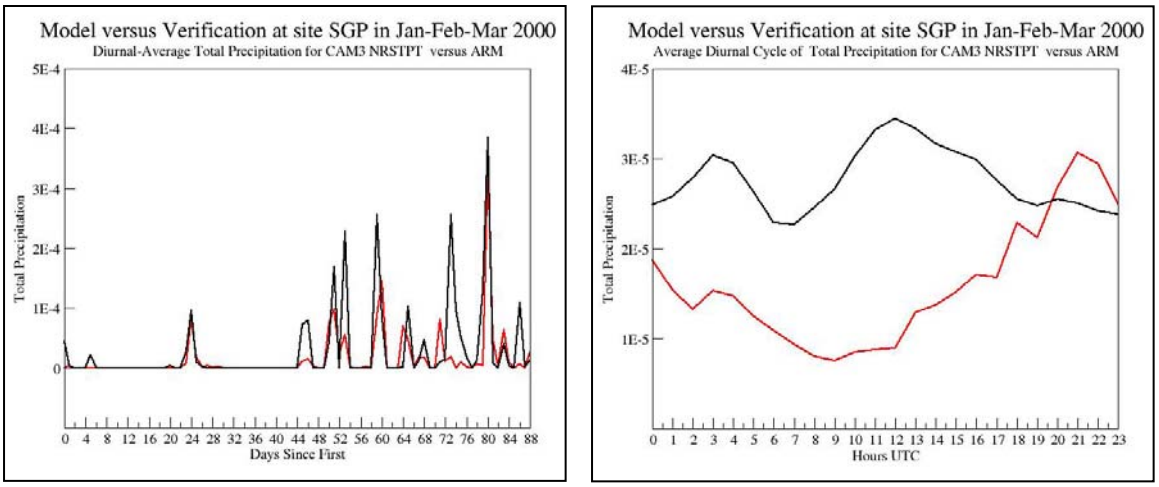
a) **b)**



During JFM, the timings of precipitation events are captured fairly well by the CAM3, but their magnitudes are generally much too low (Figure 24a). The diurnal cycle of precipitation also is qualitatively different from the observations (Figure 24b).

Figure 24: Total precipitation fluxes (units of kg m^{-2} s^{-1}) during JFM of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

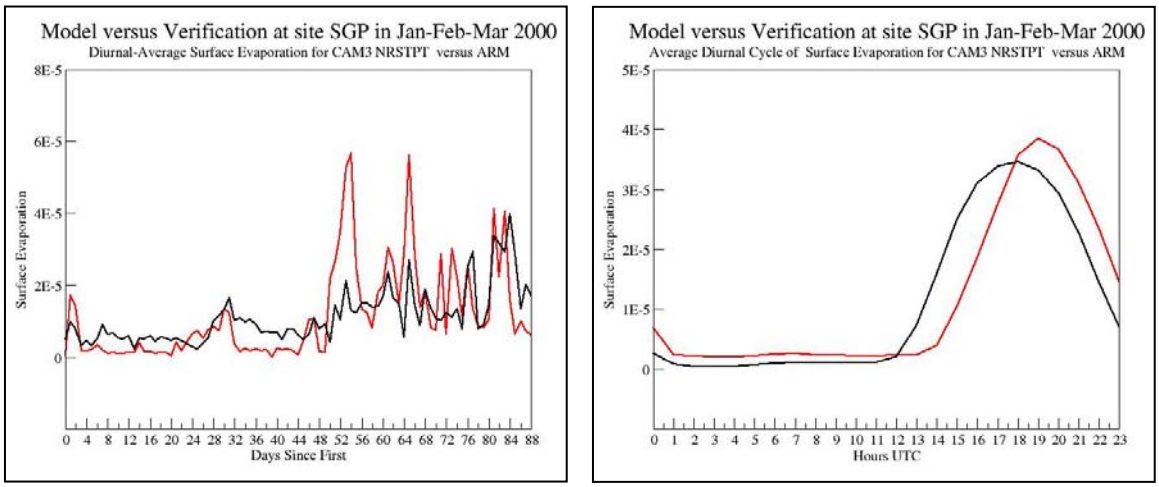
a) **b)**



Daily averages of surface evaporation move from one extreme to another: they are biased somewhat negative in the first half of the season when precipitation is infrequent, but are biased strongly positive when a chain of precipitation events are initiated after about day 50 (Figure 25a). A possible explanation for the latter strong over-prediction is that the evaporation from wet vegetation and/or bare ground surfaces may be too vigorous in this JFM season; but a poor simulation of the ABL also may be a contributing factor. The JFM-mean diurnal cycle (Figure 25b) shows a net positive bias in amplitude, but also a substantial phase lag (like the sensible heat flux, Figure 23b)

Figure 25: Surface evaporation fluxes (units of kg m^{-2} s^{-1}) during JFM of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the JFM-mean diurnal cycle.

a) **b)**



Despite the CAM3's general over-prediction of surface evaporation during JFM, the precipitable water of the entire atmosphere is under-predicted (figures not shown).

Model forecasts during the year-2000 spring season

During the April-May-June (AMJ) spring season, radiative forcings increase as the season progresses, and surface conditions fall between winter and summer extremes. AM2 and CAM3 forecasts thus occur in the context of a transitional energy forcing trend. Because year-2000 AMJ precipitation events are observed to occur more often than in JAS or JFM, moist processes at the land surface also are more fully active during this season.

AM2 forecasts during AMJ

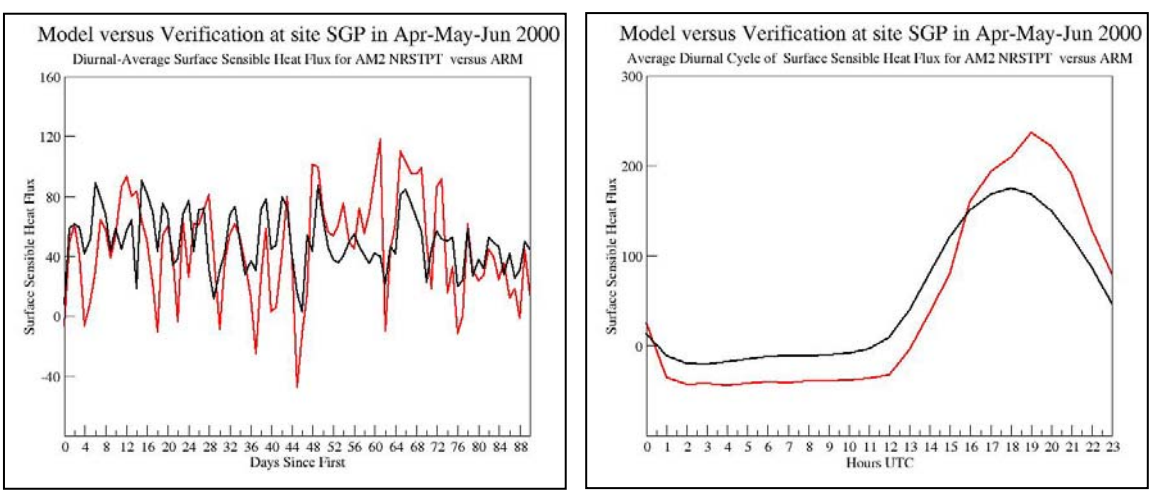
In the AM2 model during AMJ, the downward surface short-wave flux is biased positive (probably because the model-predicted cloud cover is generally too small), but the downward long-wave flux is in good agreement with the observations. Although the surface upward short-wave and long-wave fluxes are both also biased positive, the net radiation absorbed by the ground is anomalously high (figures not shown).

The surface sensible heat flux thus is generally also biased positive (regression slope **m** of hourly samples ~ 1.26, and Figure 26a), as is the amplitude of its seasonal-mean diurnal cycle during daytime hours (Figure 26b). The intra-diurnal sensible heat flux also shows a prominent phase lag relative to observations (Figure 26b).

The AMJ-mean intra-diurnal surface air temperature (not shown) is positively biased, both day and night, as well. The diurnal cycle of surface wind speed displays the same bias as noted in other seasons (e.g. Figure 6).

Figure 26: Surface sensible heat fluxes (units of W m^{-2}) during JFM of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the AMJ-mean diurnal cycle.

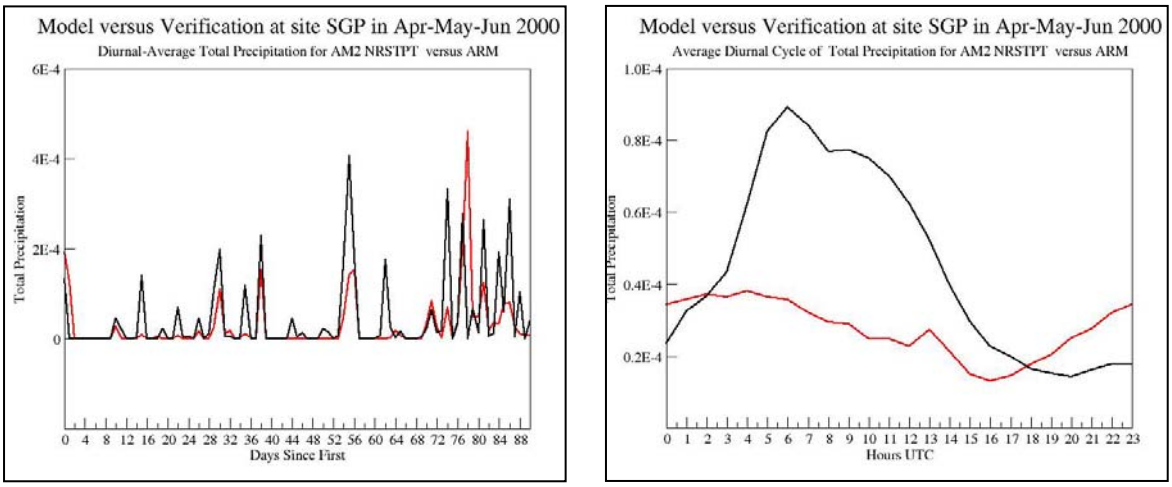
a) **b)**



The AM2 forecasts the timing of precipitation events fairly well during AMJ, but with a magnitude that is mostly too low (Figure 27a). The model precipitation errors are even more obvious for the AMJ-mean diurnal cycle, with an amplitude that is much too flat and with the prominent observed nighttime peak entirely absent (Figure 27b).

Figure 27: Total precipitation fluxes (units of kg m^{-2} s^{-1}) during AMJ of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the AMJ-mean diurnal cycle.

a) **b)**

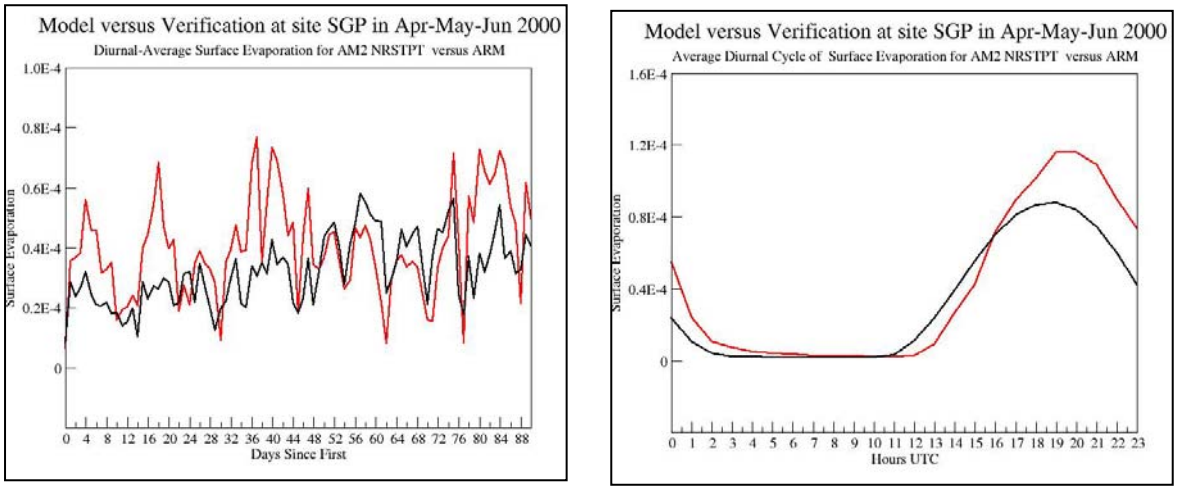


Even though the AM2 precipitation is generally too scant during AMJ, model surface evaporation is biased strongly positive in both the daily averages (Figure 28a) and in the mean diurnal cycle, where a phase lag relative to the observations is also evident (Figure 28b).

Nonetheless, the surface relative humidity is anomalously low, although the phase of its diurnal cycle compares fairly well with observations. AM2 precipitable water is biased positive, however, as in other seasons (figures not shown).

Figure 28: Surface evaporation fluxes (units of kg m^{-2} s^{-1}) during AMJ of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the AMJ-mean diurnal cycle.

a) **b)**



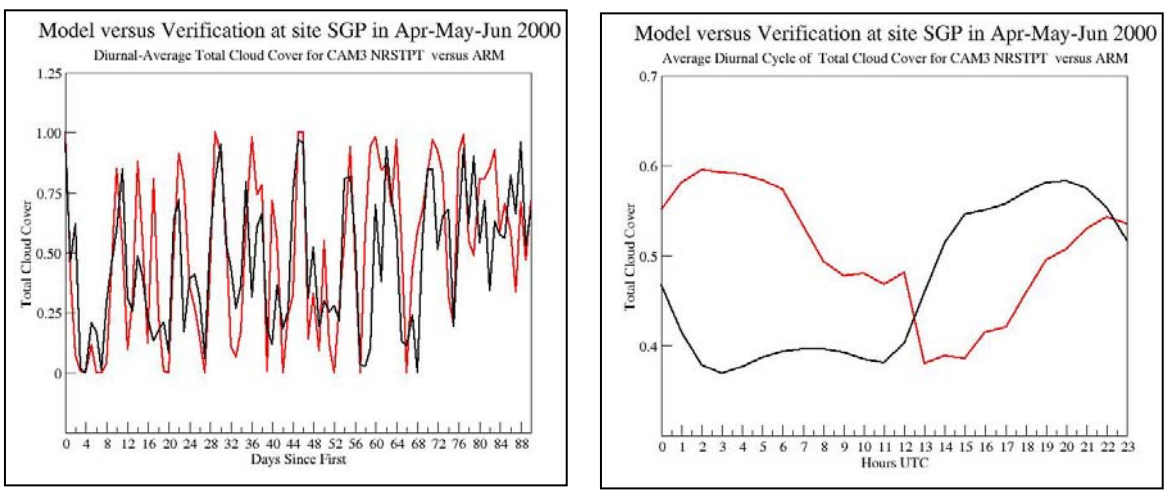
CAM3 forecasts during AMJ

In the CAM3 forecasts during AMJ, the downward surface short-wave radiation shows a sizable positive bias and a phase lag in the mean diurnal cycle as well; the net (downward minus upward) short-wave flux displays a similar pattern. The net (upward minus downward) surface long-wave flux in CAM3 is also positively biased, both day and night, with a phase lag relative to observations evident in its diurnal cycle (figures not shown)

From inspection of the vertical profile of clouds in CAM3 during AMJ (not shown), cloud amounts at high levels appear to be too large, while middle and low clouds are too sparse. The daily average total (vertically integrated) cloud amount in CAM3 is not too different from local observations (Figure 29a), but the mean model diurnal cycle (Figure 29b) is markedly out of phase: instead of peaking in late afternoon as observed, CAM3 total cloud amount is at a maximum in the local evening hours and is at a minimum in early afternoon. This pattern is consistent with the positively biased diurnal cycle of downward surface short-wave radiation noted above.

Figure 29: Total cloud cover fractions during AMJ of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the AMJ-mean diurnal cycle.

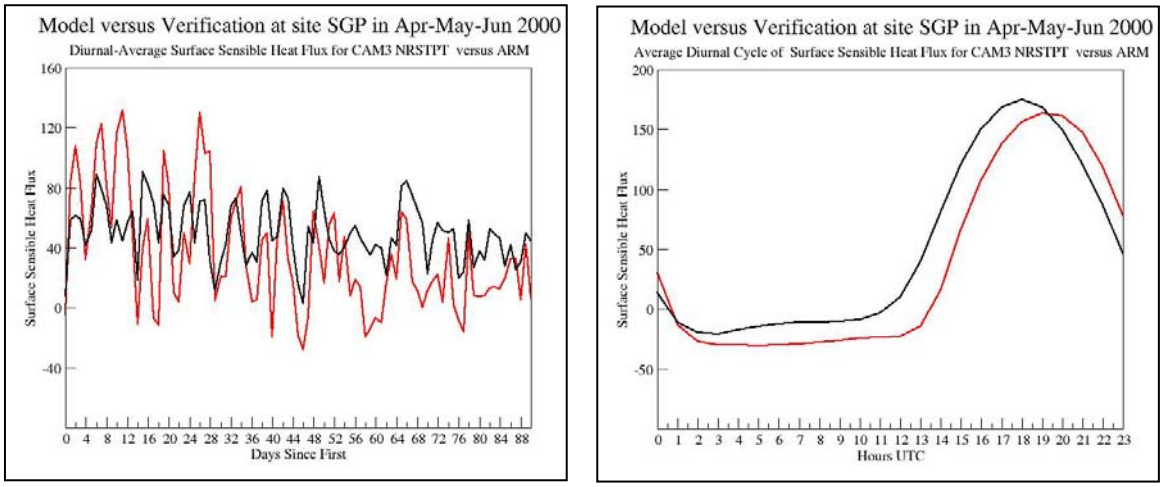
a) **b)**



The daily averages of the CAM3 surface sensible heat flux and the amplitude of its AMJ-mean diurnal cycle are biased negative, however (Figure 30). The diurnal cycle (Figure 30b) also shows a pronounced phase lag, implying over-prediction of the ground heat storage.

Figure 30: Surface sensible heat fluxes (units of W m^{-2}) during AMJ of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the AMJ-mean diurnal cycle.

a) **b)**

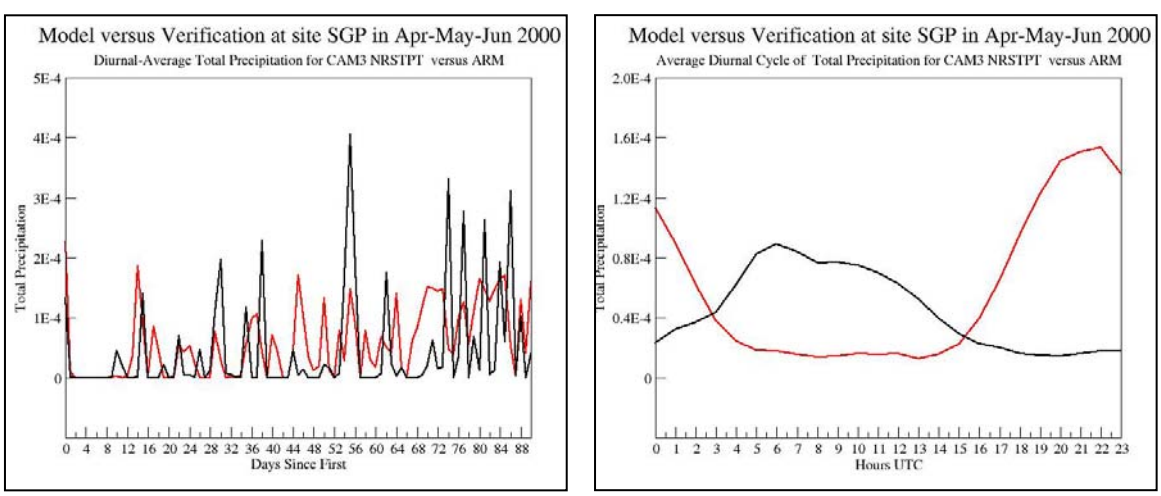


The diurnal cycle of surface air temperature is similarly biased and phase-shifted, while that of the lowest-layer wind displays the same pattern as in other seasons (i.e. anomalously high nighttime winds, and low daytime winds--see Figure 6).

The CAM3's predicted timings of daily average precipitation during AMJ agree somewhat with observations, but forecasts of precipitation magnitudes are much less so: excessive precipitation is predicted for several observed events, while smeared-out rather than sharply punctuated peaks occur for many others (Figure 31a). The CAM3 diurnal cycle of precipitation displays an obvious positive bias, with a peak value occurring in mid-afternoon rather than around local midnight as observed (Figure 31b).

Figure 31: Total precipitation fluxes (units of kg m^{-2} s^{-1}) during AMJ of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the AMJ-mean diurnal cycle.

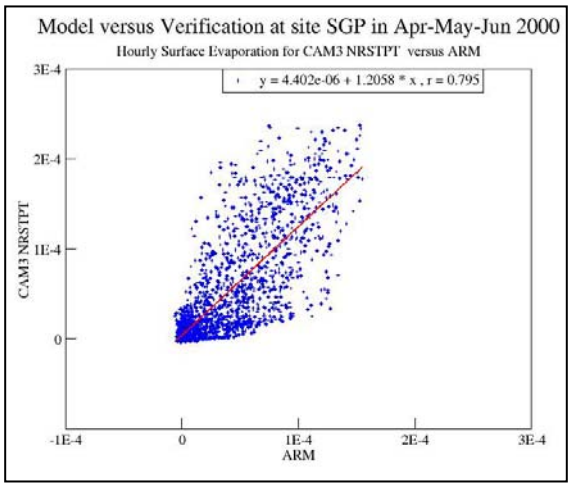
a) **b)**



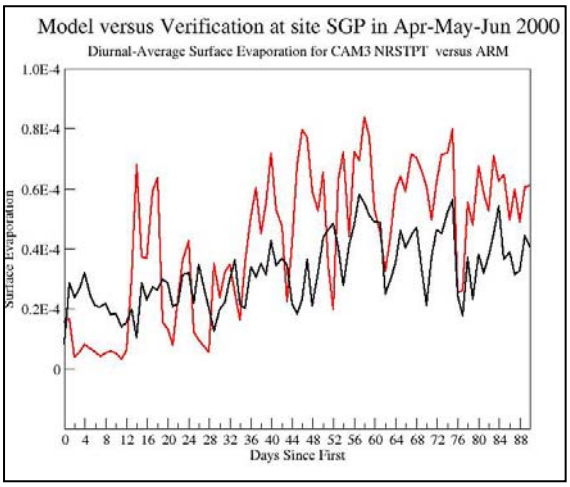
CAM3 surface evaporation (Figure 32) also displays an obvious positive bias (regression slope **m** of hourly samples is ~ 1.2) at all time scales, with a phase lag evident in the diurnal cycle. Nevertheless, precipitable water also is substantially under-predicted (figures not shown)

Figure 32: Surface evaporation fluxes (units of kg m^{-2} s^{-1}) during AMJ of the year 2000 are depicted in a) as a scatter of hourly samples of CAM3 forecasts vs. ARM observations. Surface evaporation fluxes for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the AMJ-mean diurnal cycle.

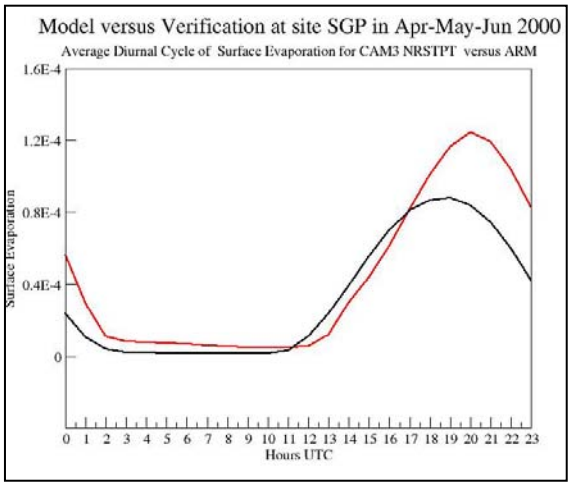
a)



b)



c)



Model forecasts during the year-2000 autumn season

During the October-November-December (OND) autumn season, the AM2 and CAM3 forecasts are implemented in the context of steadily decreasing seasonal radiative forcings. As in the AMJ spring season, year-2000 OND precipitation events occur frequently, so that hydrological forcing tends to play an increasingly important role relative to radiative forcing as the season progresses.

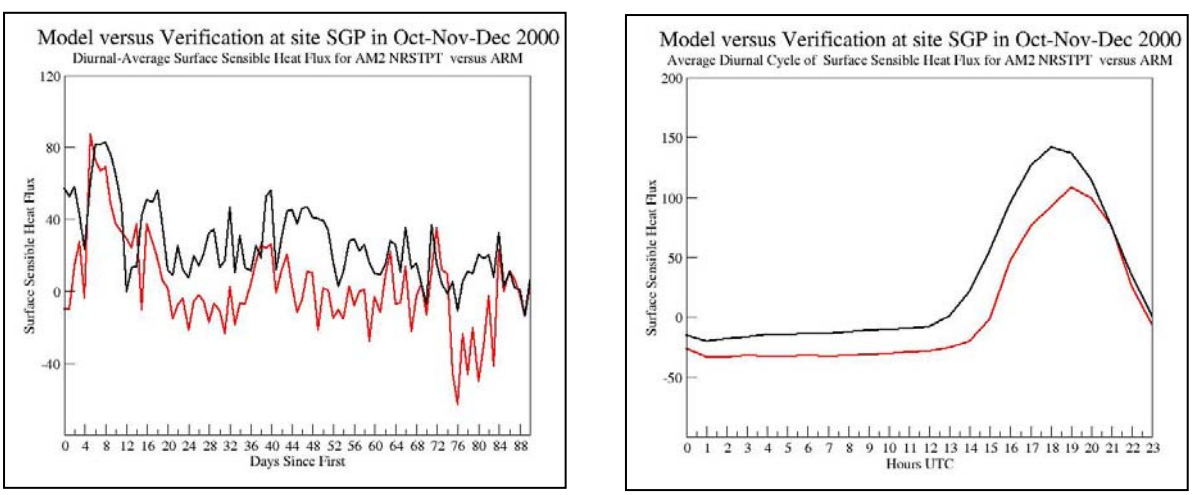
AM2 forecasts during OND

During OND, the AM2 displays positively biased downward short-wave and negatively biased downward long-wave radiative fluxes during daylight hours (figures not shown). These results suggest that daytime cloud amounts may be under-predicted. (Total cloud amount is not output by AM2, so this premise cannot be checked directly.) Biases are generally less evident in the upward surface short-wave and long-wave fluxes (figures not shown).

Despite the over-prediction of surface short-wave radiation, the surface sensible heat flux shows a sizeable negative bias in both daily averages and in its OND-mean diurnal cycle (Figure 33), as well as at hourly time scales (regression slope **m** ~ 0.82). This negative bias in sensible heat flux seems to be an outcome of a very large positive bias in surface evaporation (discussed below).

Figure 33: Surface sensible heat fluxes (units of W m^{-2}) during OND of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the OND-mean diurnal cycle.

a) **b)**

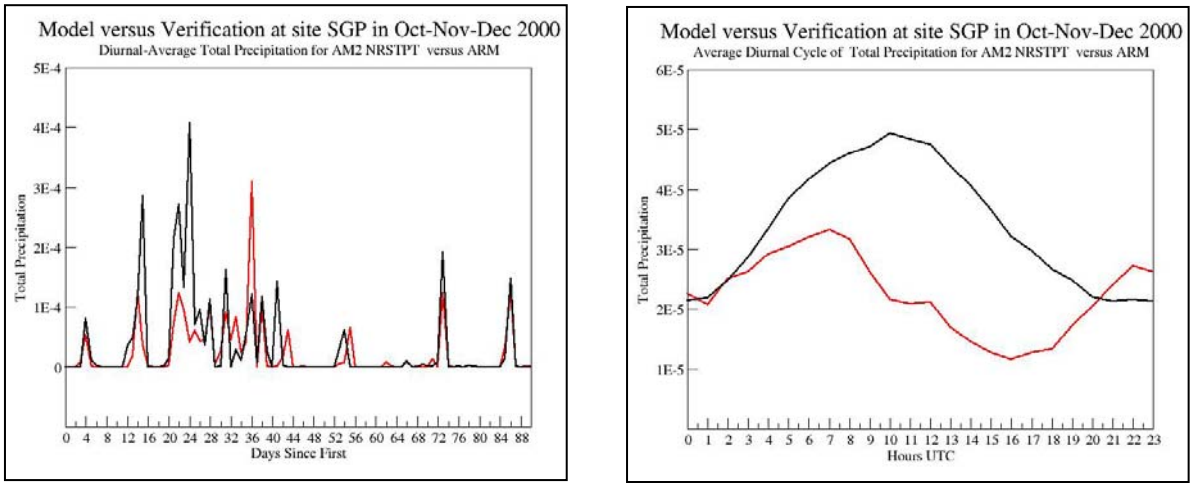


Surface air temperatures are well-predicted, however, with only a small positive bias at night, and a negative bias during the daylight hours (figure not shown). The surface winds display an intra-diurnal bias similar to that seen in other seasons (e.g. Figure 6).

The AM2 forecasts the timing of OND daily precipitation events fairly well (Figure 34a), but their magnitudes are generally under-predicted. The depressed precipitation magnitude is even more apparent in the model's OND-mean diurnal cycle (Figure 34b), and peak precipitation occurs just after midnight, several hours before the observed intra-diurnal peak.

Figure 34: Total precipitation fluxes (units of kg m^{-2} s^{-1}) during OND of the year 2000 for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the OND-mean diurnal cycle.

a) **b)**

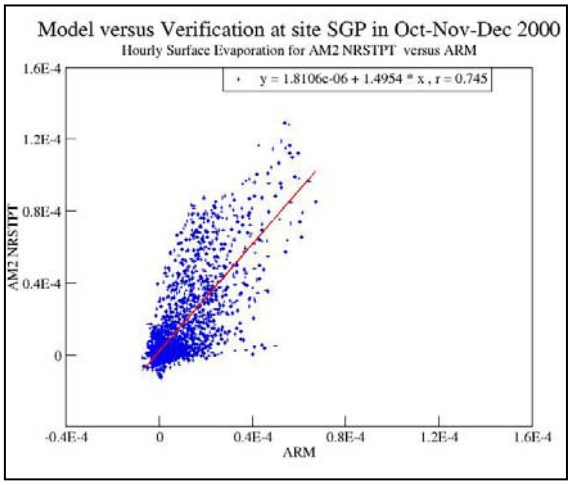


Surface evaporation (Figure 35) nonetheless displays a very large positive bias at all time scales, thereby depressing the need for energy transfer by the surface sensible heat flux (Figure 33). These results are symptomatic of an incorrect partitioning of the turbulent fluxes by the AM2's land model.

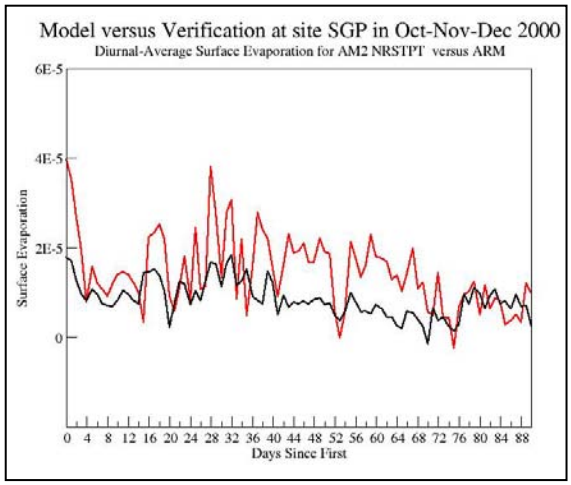
AM2 daily average surface relative humidity also generally shows a large positive bias, but the amplitude of the OND-mean diurnal cycle is too flat compared with that observed. Precipitable water is also substantially over-predicted (figures not shown).

Figure 35: Surface evaporation fluxes (units of kg $m^{-2} s^{-1}$) during OND of the year 2000 are depicted in a) as a scatter of hourly samples of AM2 forecasts vs. ARM observations. Surface evaporation fluxes for ARM observations (in black) vs. AM2 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the OND-mean diurnal cycle.

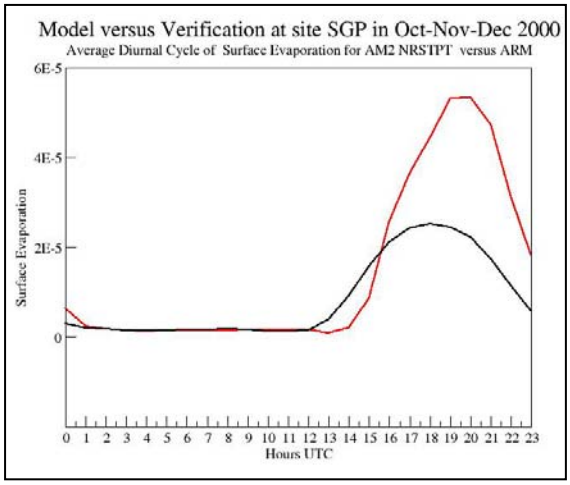
a)



b)



c)

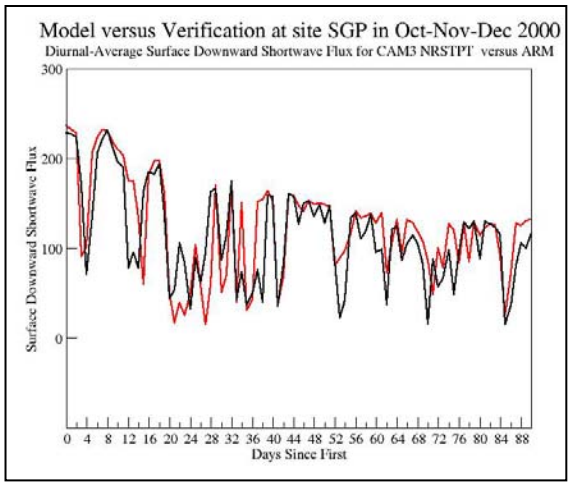


CAM3 forecasts during OND

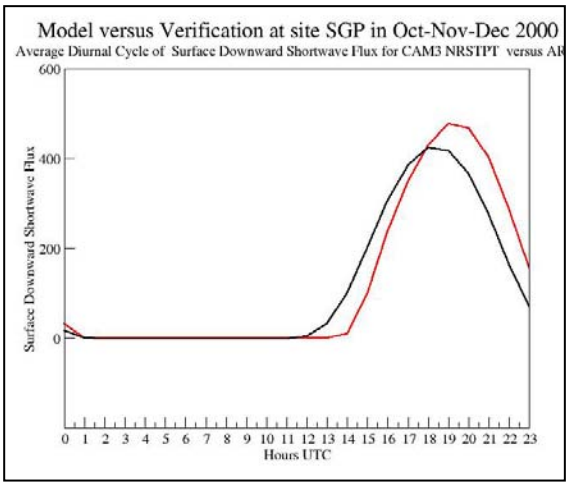
The daily average downward surface short-wave flux during OND is in fairly good agreement with observations (Figure 36a), but its mean diurnal cycle shows substantial positive bias and phase lag relative to the ARM observations (Figure 36b). Similar bias and phase errors are evident in the net surface short-wave radiation.

Figure 36: Surface downward shortwave fluxes (units of W m^{-2}) during OND of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the OND-mean diurnal cycle. In c) the corresponding observed (in black) and model (in red) diurnal cycle of total cloud cover fractions are shown.

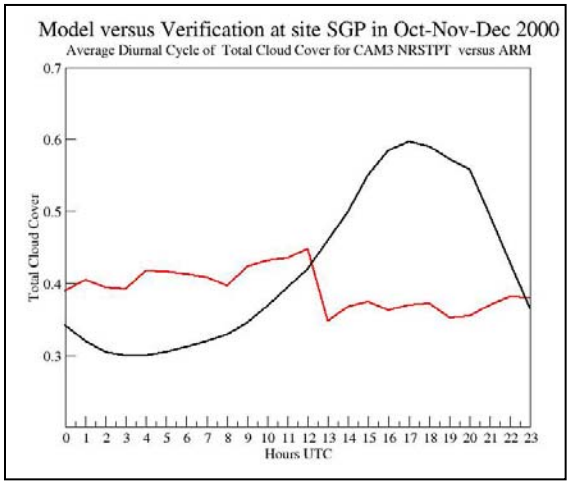
a)



b)



c)



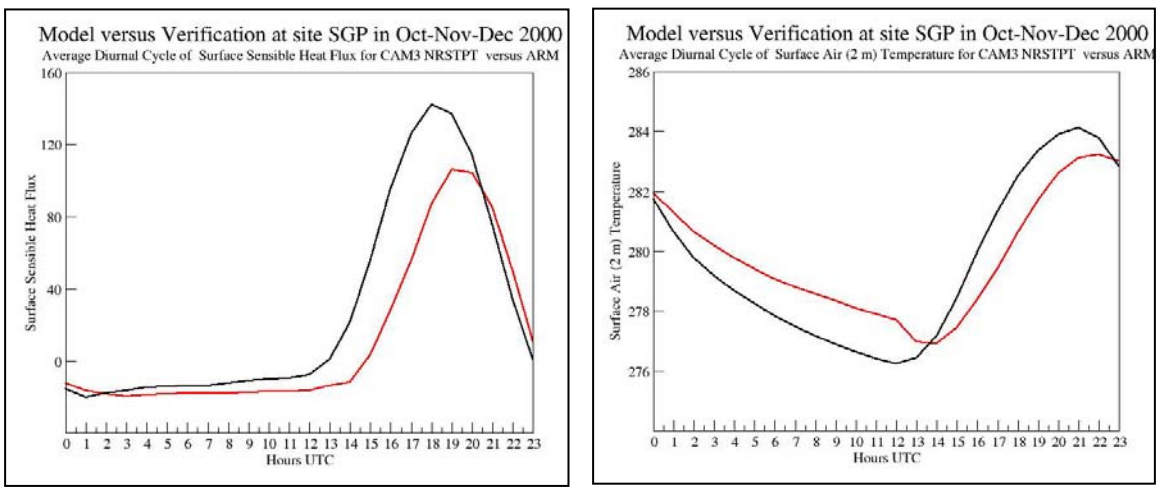
The net long-wave radiation (not shown) also displays substantial positive biases in both its daily averages and OND-mean diurnal cycle (especially at night for the latter), implying too much emission from the surface and/or too little downward long-wave flux from atmospheric moisture and clouds.

These bias and phase errors in the CAM3 radiative fluxes are probably related to model anomalies in the diurnal cycle of total cloud cover, whose amplitude is markedly depressed, and whose phase is anti-correlated with that of the observations (Figure 36c). The vertical profile of cloud amounts in CAM3 (not shown) implies that total cloud cover is insufficient mostly because model clouds are too sparse at low and middle vertical levels.

Despite the excessive downward surface shortwave radiation, daily averages of the surface sensible heat flux (not shown) are negatively biased. This also is the case throughout the entire OND-mean diurnal cycle (Figure 37a), with an excessive sensible heat flux being directed downward from screen level to ground at night. This pattern reflects an erroneous partitioning of the net available surface energy into sensible vs. latent heating (see discussion of surface evaporation below). The surface air temperature (Figure 37b) is too low during daylight hours but is too high at night--consistent with the excessive downward nocturnal sensible heat flux.

Figure 37: OND-mean diurnal cycle of surface sensible heat fluxes (units of W m^{-2}) in the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a), and in b) is displayed the corresponding OND-mean diurnal cycles of surface air temperatures.

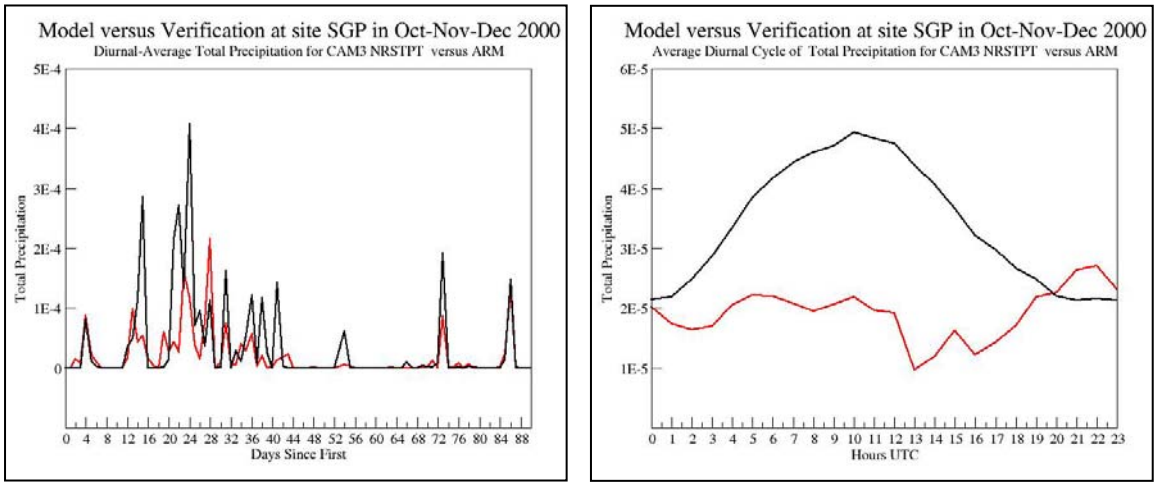
a) **b)**



During OND, the CAM3 timings of precipitation events are fairly good, but the amounts are generally much too low (Figure 38 a). This deficiency is even more apparent in the modeled OND-mean diurnal cycle of precipitation (Figure 38b), which completely misses the sharp early-morning peak in the observations.

Figure 38: Total precipitation fluxes (units of kg m^{-2} s^{-1}) during OND of the year 2000 for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in a) as time series of daily averages, and in b) as the OND-mean diurnal cycle.

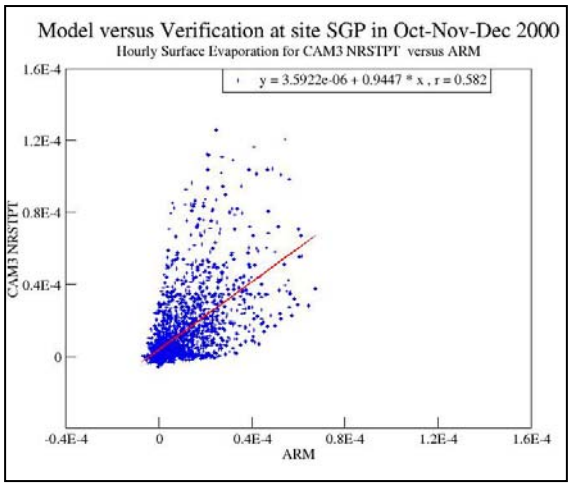
a) **b)**



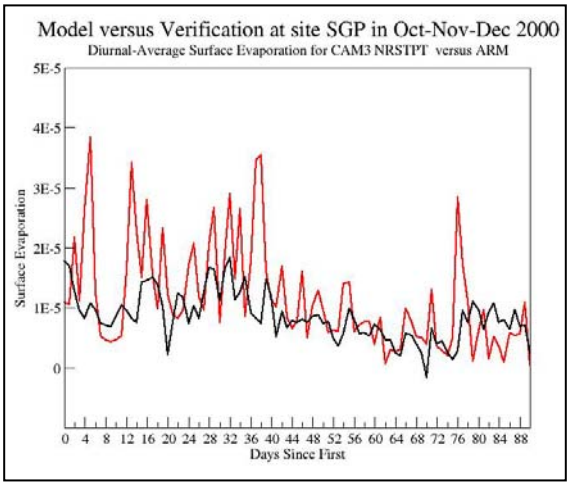
Despite the generally sparse CAM3 precipitation, the model surface evaporation shows large positive biases (more so during the day, but also at night) relative to local observations at all time scales (Figure 39). However, the precipitable water displays its usual negative bias for this model (not shown), implying that the excessive surface moisture flux does not impact higher levels of the CAM3 atmosphere.

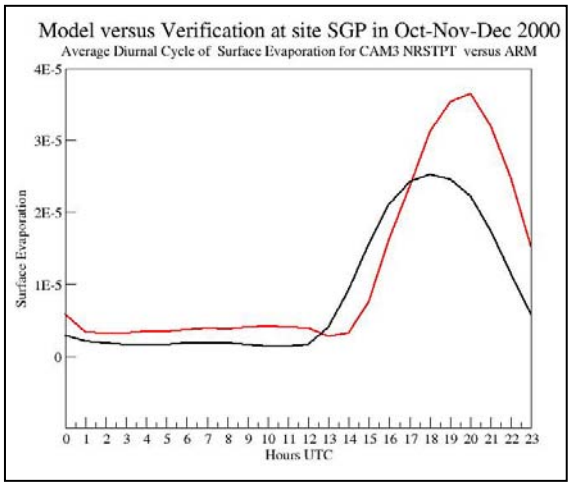
Figure 39: Surface evaporation fluxes (units of kg m^{-2} s^{-1}) during OND of the year 2000 are depicted in a) as a scatter of hourly samples of CAM3 forecasts vs. ARM observations. Surface evaporation fluxes for ARM observations (in black) vs. CAM3 forecasts (in red) are depicted in b) as time series of daily averages, and in c) as the OND-mean diurnal cycle.

a)



b)





Implications of systematic forecast errors for model schemes

The AM2 and CAM3 year-2000 forecasts of surface variables at the ARM SGP site display a number of similar errors in forcing variables. For example, both models generally show positively biased downward surface short-wave radiative fluxes, a persistent problem in many GCMs (e.g. Garratt et al. 1993). This error appears to be tied to a general under-prediction of total cloud cover. In CAM3 where it can be tracked more directly, the diurnal cycle of total cloud cover also deviates qualitatively from the observations. Further attention to cloud formation schemes is therefore clearly still needed.

Prediction of hydrological forcing via precipitation is also a serious weakness of both models. While the AM2 forecasts the timings of precipitation events with some success, it generally under-predicts precipitation magnitude. The CAM3 is more erratic in predicting the timings of precipitation events, and it also shows a marked tendency to produce "convective drizzle" in the summer season. Moreover, both models qualitatively fail to simulate the observed seasonal-mean diurnal cycle of precipitation. There is thus an obvious need for improvement of the models' simulation of precipitation formation. The more egregious failings in precipitation prediction during the summer season suggest that especial attention to the representation of convection is called for in both models.

In analyzing the surface response to these forcings, the evaporative flux, as the key "bridge" variable between the surface energy and moisture budgets, is a telling indicator of problematical model behaviors. For example, over-prediction of surface evaporation (and its common corollary, under-prediction of surface sensible heat flux) appears to be a particular problem in the AM2, especially considering this model's general under-prediction of precipitation. Although the excessive surface evaporation may partly be a result of the positive biases downward surface short-wave radiation, this is unlikely to be the entire explanation: for example, large positive biases in surface evaporation occur in the winter season, when the land surface experiences minimal radiative stress. Moreover, the AM2's land model is a modified "bucket" scheme (Milly and Shmakin 2002) that is known to be susceptible to such a problem (reference: PILPS studies). Given the systematic errors in the diurnal cycle of model surface winds, there also may be a need to consider the details of the model bulk transfer schemes that regulate the turbulent transports. It should be noted, however, that correcting biases in the surface evaporation may also entail making changes in the representation of runoff, a variable not available for the present study. While the CAM3 employs the more complex Community Land Model (CLM, Oleson et al. 2004) that seems to exhibit a less acute problem of this type, the large phase lags seen in the CAM3's seasonal-mean diurnal cycles of surface evaporation and sensible heat flux suggest that the CLM's representation of ground heat storage may be faulty. Thus, there is an apparent need to consider making further modifications in the land models of both AM2 and CAM3.

References

Boyle, J.S., D. Williamson, R. Cederwall, M. Fiorino, J. Hnilo, J. Olson, T. Phillips, G. Potter, and S. Xie, 2005: Diagnosis of Community Atmospheric Model 2 (CAM2) in numerical weather forecast configuration at Atmospheric Radiation Measurement sites. *Journal of Geophysical Research*, **110**, D15516, doi: 10.1029/2004JD005109.

Garratt, J.R., P.B. Krumal, and E.A. Kowalczyk, 1993: The surface energy budget at local and regional scales—A comparison of general circulation model result with observations. *J. Climate*, **6**, 1090-1109.

Jiang, X., N.-C. Lau, and S. A. Klein, 2006: Role of eastward propagating convection systems in the diurnal cycle and seasonal mean of summertime rainfall over the U.S. Great Plains. *Geophys. Res. Lett*., **33**, doi:10.1029/2006GL027022.

Klein, S.A., X. Jiang, J. Boyle, S. Malyshev, and S. Xie, 2006: Diagnosis of the summertime warm and dry bias over the U.S. Southern Great Plains in the GFDL climate model using a weather forecasting approach. *Geophys. Res. Lett.*, **33**, L18805, doi: 10.1029/2006GL027567.

Milly, P.C.D., and A.B. Shmakin, 2002: global modeling of land water and energy balances, Part I: The Land Dynamics (LaD) model. *J. Hydrometeor*., **3**, 283-299.

Oleson, K.W., and Coauthors, 2004: Technical description of the Community Land Model (CLM). Tech. Rep. NCAR/TN-461+STR, National Center for Atmospheric Research, Boulder, CO, 174 pp.

Svensson, G., and B. Holtslag, 2007 : The diurnal cycle—GABLS Second Intercomparison Project. *GEWEX News*, Vol. 17, No. 1, 9-10.

Viterbo, P., A. Beljaars, J-F. Mahfouf, and J. Teixeira, 1999: The representation of soil moisture freezing and its impact on the stable boundary layer. *Q. J. Roy. Meteor. Soc*., **125**, 2401-2426.

Williamson, D. L., J. Boyle, R. Cederwall, M. Fiorino, J. Hnilo, J. Olson, T. Phillips, G. Potter, and S. Xie, 2005: Moisture and temperature balances at the Atmospheric Radiation Measurement Southern Great Plains Site in forecasts with the CAM2. *J. Geophys. Re s*., **110**, D15S16,doi:10.1029/2004JD005109.

Williamson, D., and J. Olson, 2007: A comparison of forecast errors in CAM2 and CAM3 at the ARM Southern Great Plains Site. *J. Climate* (in press).

Xie, S., M. Zhang, J.S. Boyle, R.T. Cederwall, G.L. Potter, and W. Lin, 2004: Impact of a revised convective triggering mechanism on Community Atmosphere Model, Version 2, simulations: Results from short-range weather forecasts. *J. Geophys. Res.*, **109**, D14102, doi:10.1029/2004JD004692.