Building a Comprehensive Evaluation Framework for VIC Model in NLDAS Testbed

Youlong Xia[1], Sujay V. Kumar[2], David M. Mocko[2], Shugong Wang[2], Christa D. Peters-Lidard[3], Helin Wei[1], and Michael B. Ek[1]

[1]Environmental Modeling Center, National Centers for Environmental Prediction, NOAA, College Park, Maryland

[2]Hydrological Sciences Laboratory, GSFC/NASA, Greenbelt, Maryland

[3]Hydrosphere, Biosphere, and Geophysics in the Earth Sciences Division, GSFC/NASA, Greenbelt, Maryland

Time scale: monthly

Reference Data:
Energy: NASA/GEWEX Radiation, MTE Energy Flux
Water: MTE and GLEAM ET, USGS Q
State: SM, ST, SNODAS SWE

Model: VIC403, VIC405, VIC412

Statistics Metrics:
Bias/Rbias – model systematic error
RMSE – model overall error
AC – simulation skill (temporal variation)
Taylor S/NSE – general combined score

dNSE~-dRMSEx(RMSE1+RMSE2)/σ^2
Rbias (student t-test) and dAC (Zou approach)
–significance test at 95% confidence level)
dRMSE=RMSE2-RMSE1, dS=S2-S1, dAC=AC2-AC1

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods, 12,* 399-413. – depended dataset

VIC Releases 4.0.0 - 4.0.6 all contain the same basic feature set:
Multiple land cover tiles per grid cell
Single soil column per grid cell, consisting of multiple soil layers
Water balance and energy balance modes
Optional finite-difference solution to soil temperature profile
Modeling of both the surface snow pack and the snow intercepted by the canopy
Simulation of seasonally-frozen soil
Multiple snow/elevation bands within a grid cell

↓

VIC Releases 4.1.0 and 4.1.1 contain several new features:
Canopy temperature no longer assumed to equal air temperature; temperature of canopy snow found by iterating to close the canopy energy balance.
Optional iteration between the canopy energy solution and the surface energy solution. This behavior is controlled by the CLOSE_ENERGY parameter.
Parameterization of spatial variability of snow cover and frozen soil
Simulation of effects of snowmelt for partial snow cover
Modeling of lakes and wetlands
Modeling of blowing snow sublimation
Modeling of permafrost
Some bugs fix

↓

VIC Releases 4.1.2 contain new updates:
Added computation of water table position
Extended the computation of soil temperatures, ice contents, and ground fluxes to all modes of model operation.
Some bugs fix

Figure 1

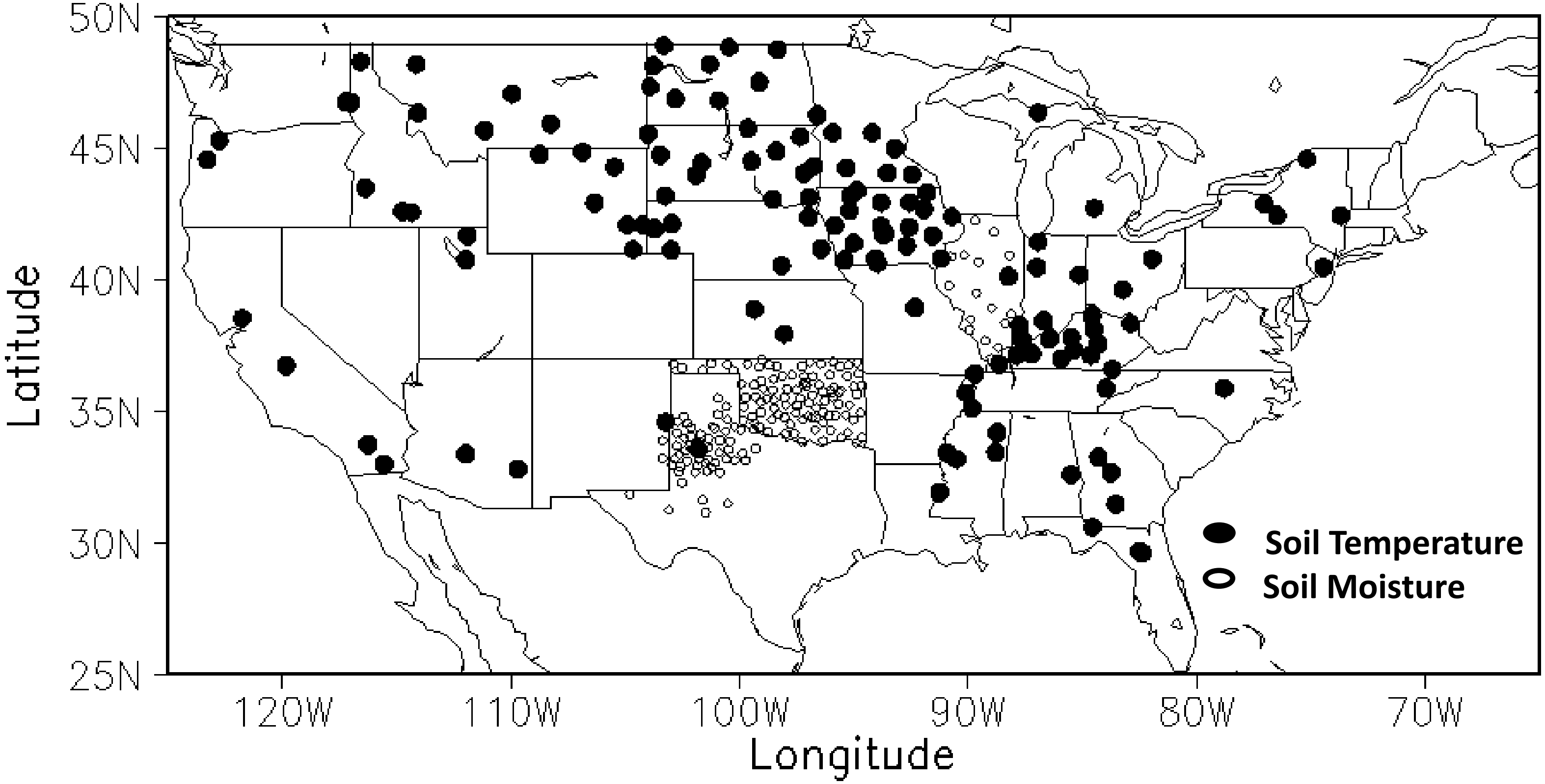


Figure 2: 137 US cooperative stations measuring soil temperature , Illinois, Oklahoma Mesonet, and West Texas Mesonet soil moisture measurement sites

Streamflow calibration reduces bias, which mostly comes from calibrated parameter (also see Figure 12)

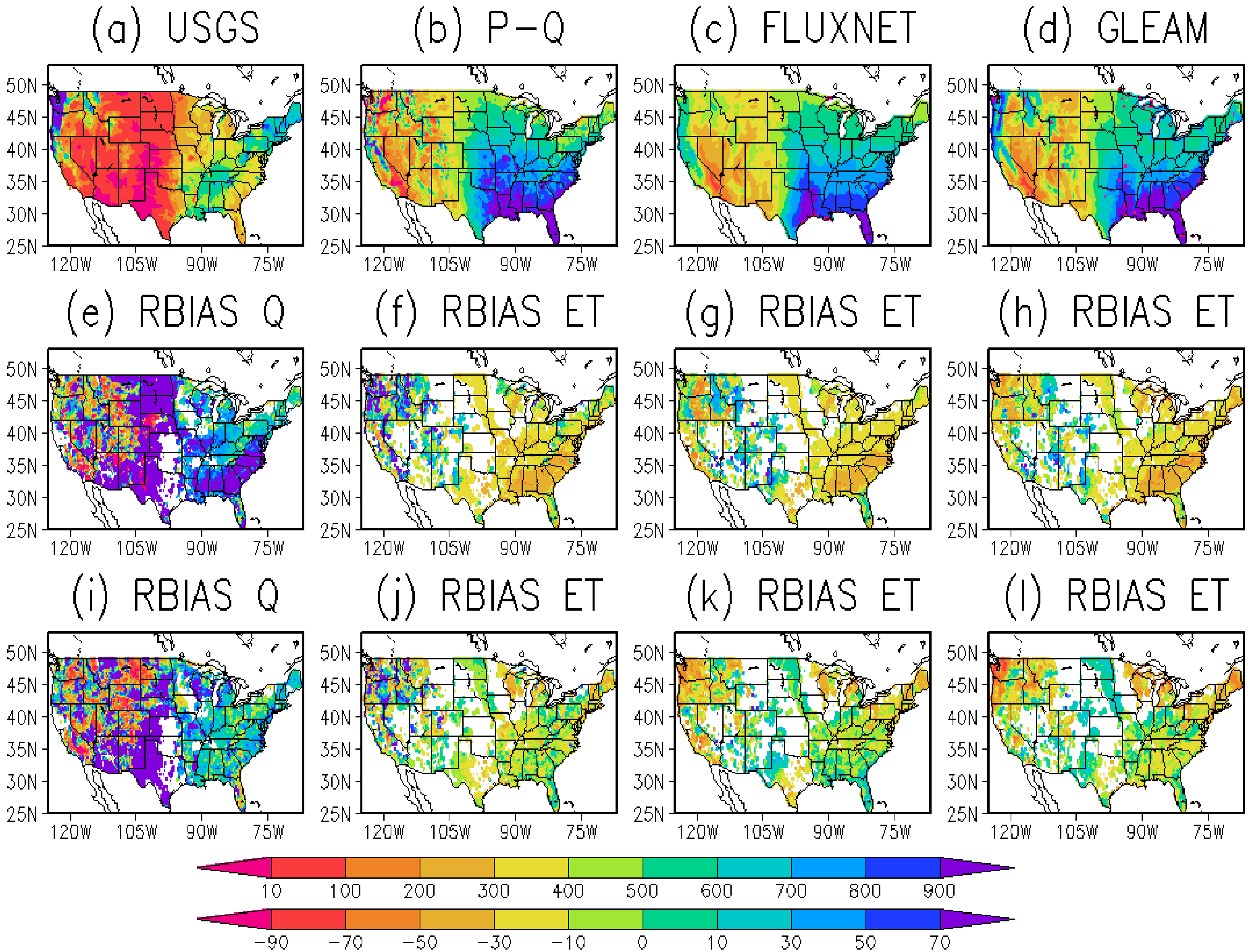


Figure 3: Mean annual Q, ET and RBIAS (RBIAS is significant at 95% confidence level)

ET (LH)) Rbias reduction leads to SH Rbias reduction

VIC412 Rbias sign changes in some region due to canopy temperature calculation treatment (See Figure 12)

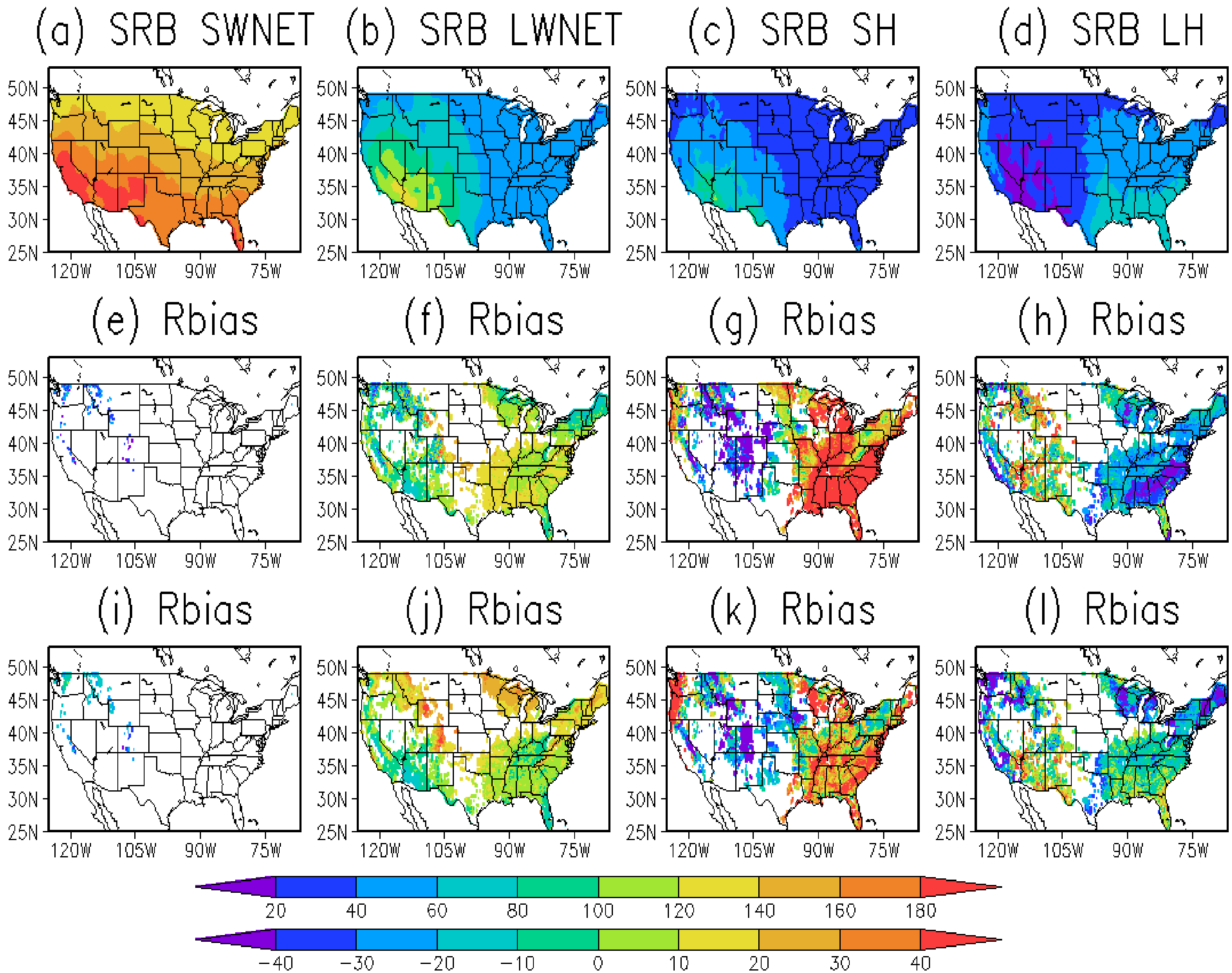


Figure 4: Mean net shortwave, net longwave, SH, LH and Rbias (insignificant Rbias is masked out)

Suggesting snow process upgrade has positive impact for many regions

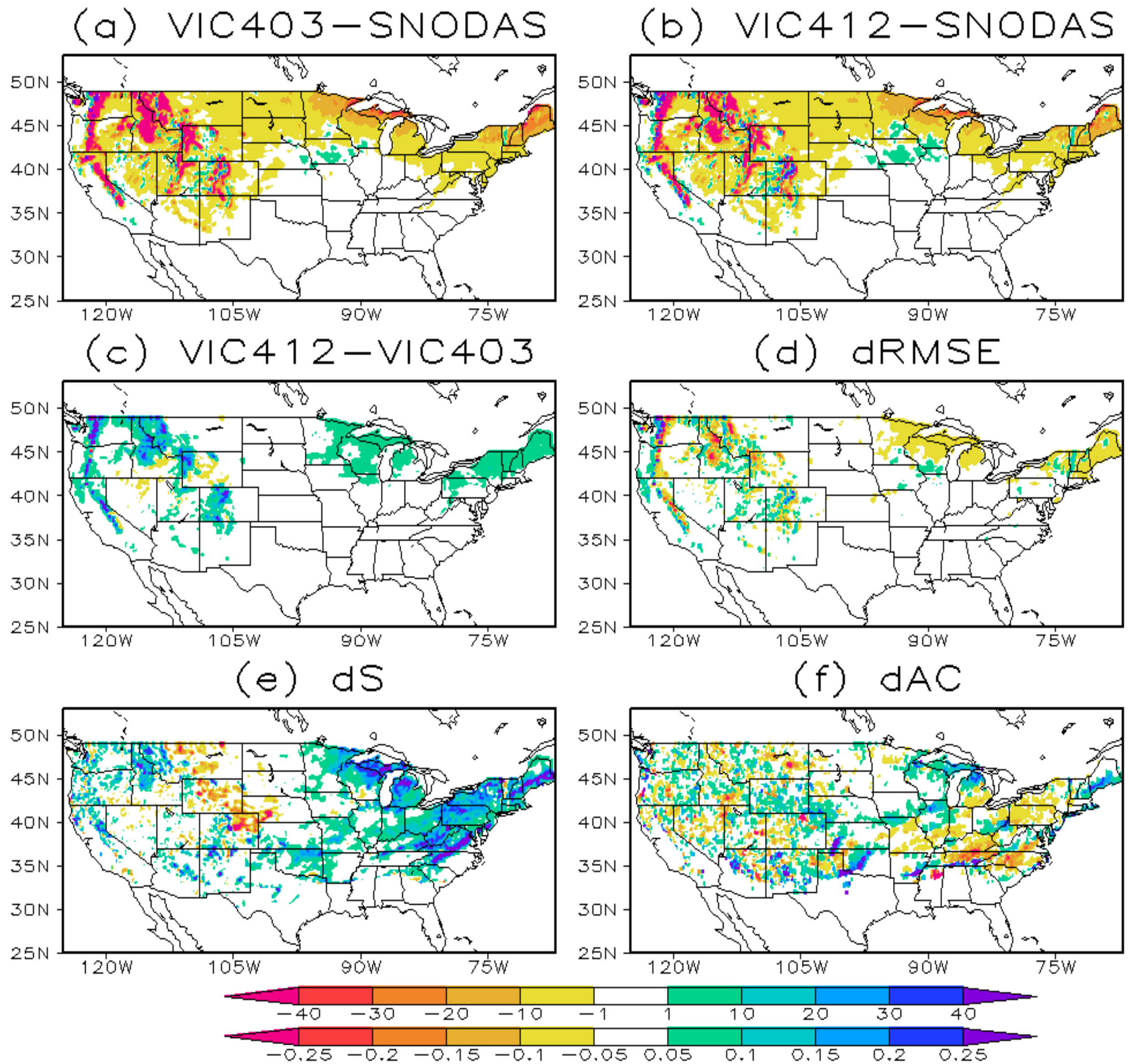


Figure 5: Bias (a,b), model difference (c), dRMSE, dS, and dAC when SNODAS SWE is used as a reference

Most improvement occurs in warm season

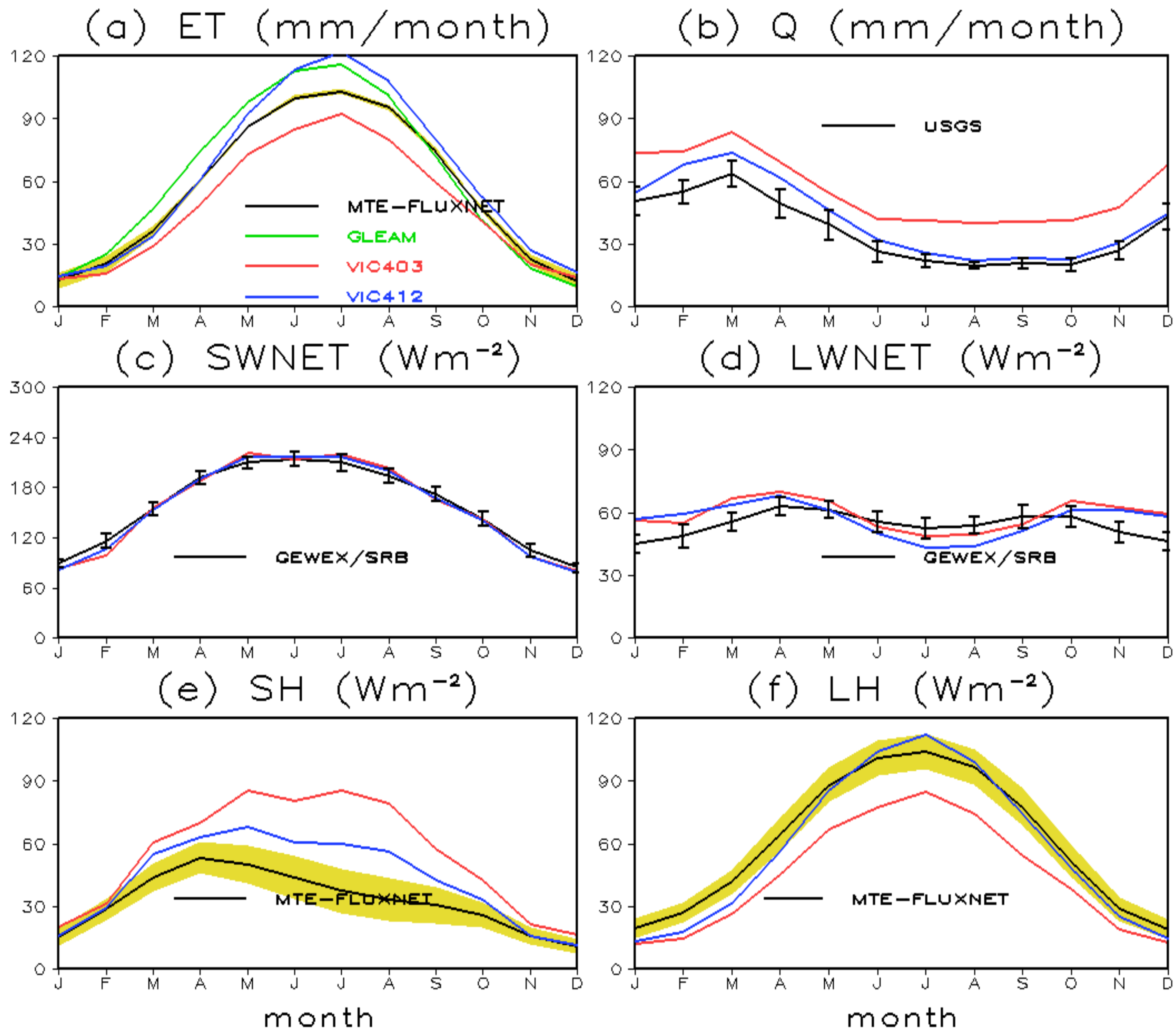


Figure 6: Mean seasonal cycles for (a) ET, (b) Q, (c) SWNET, (d) LWNET, (e) SH, and (f) LH over the southeastern US.

Small impact on soil temperature.

Mixed results for soil moisture seasonal cycles

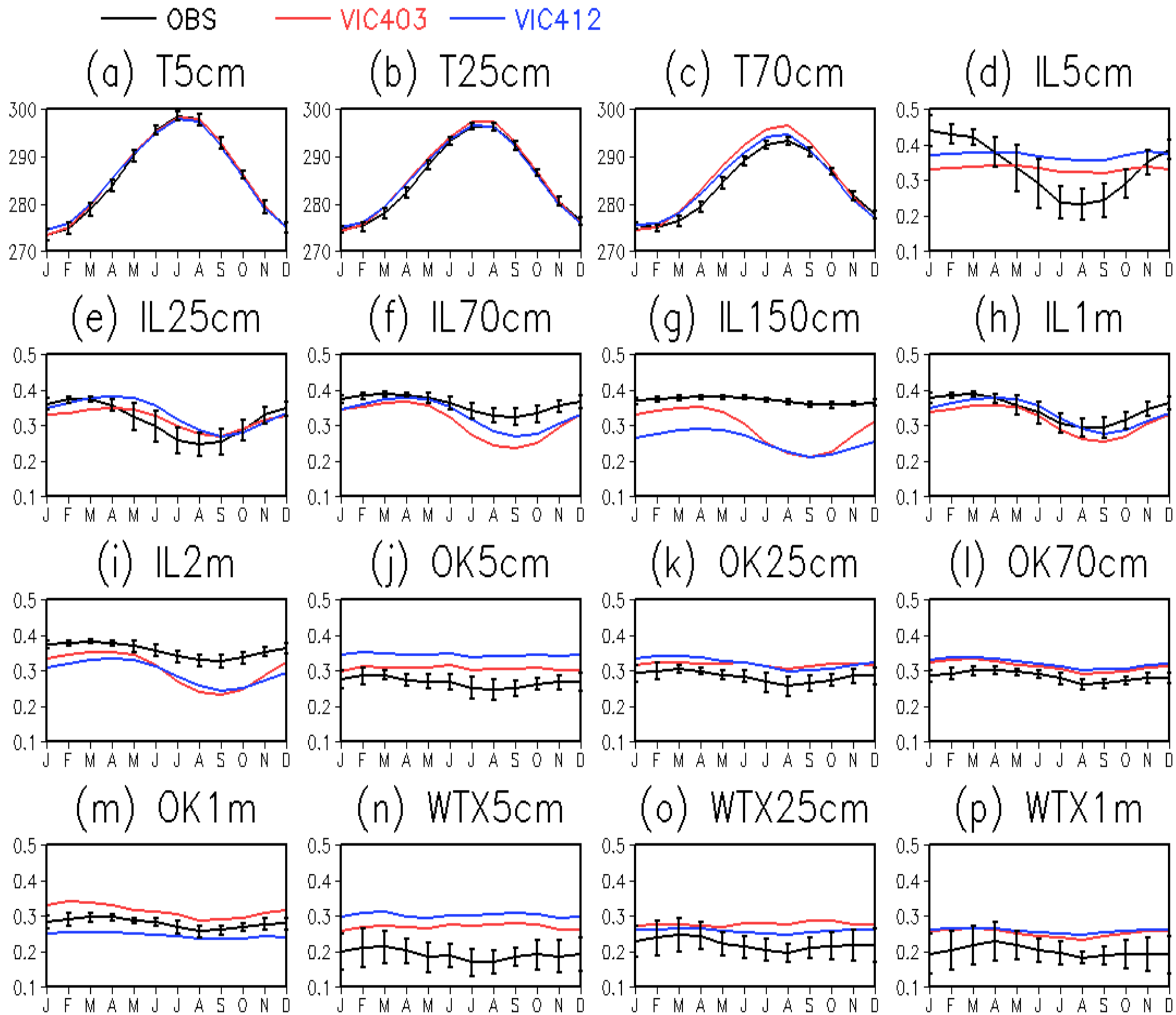


Figure 7: Mean seasonal cycle for soil temperature and moisture

Q RMSE reduction is consistent with NSE calibration (Troy et al., 2008). How this does not suggests S improvement (see Figure S3), suggesting that which metrics should be used for calibration and evaluation is not trivial.

The evaluation with two ET products generate some completely different results in some regions, suggesting that using a single ET product may be questionable.

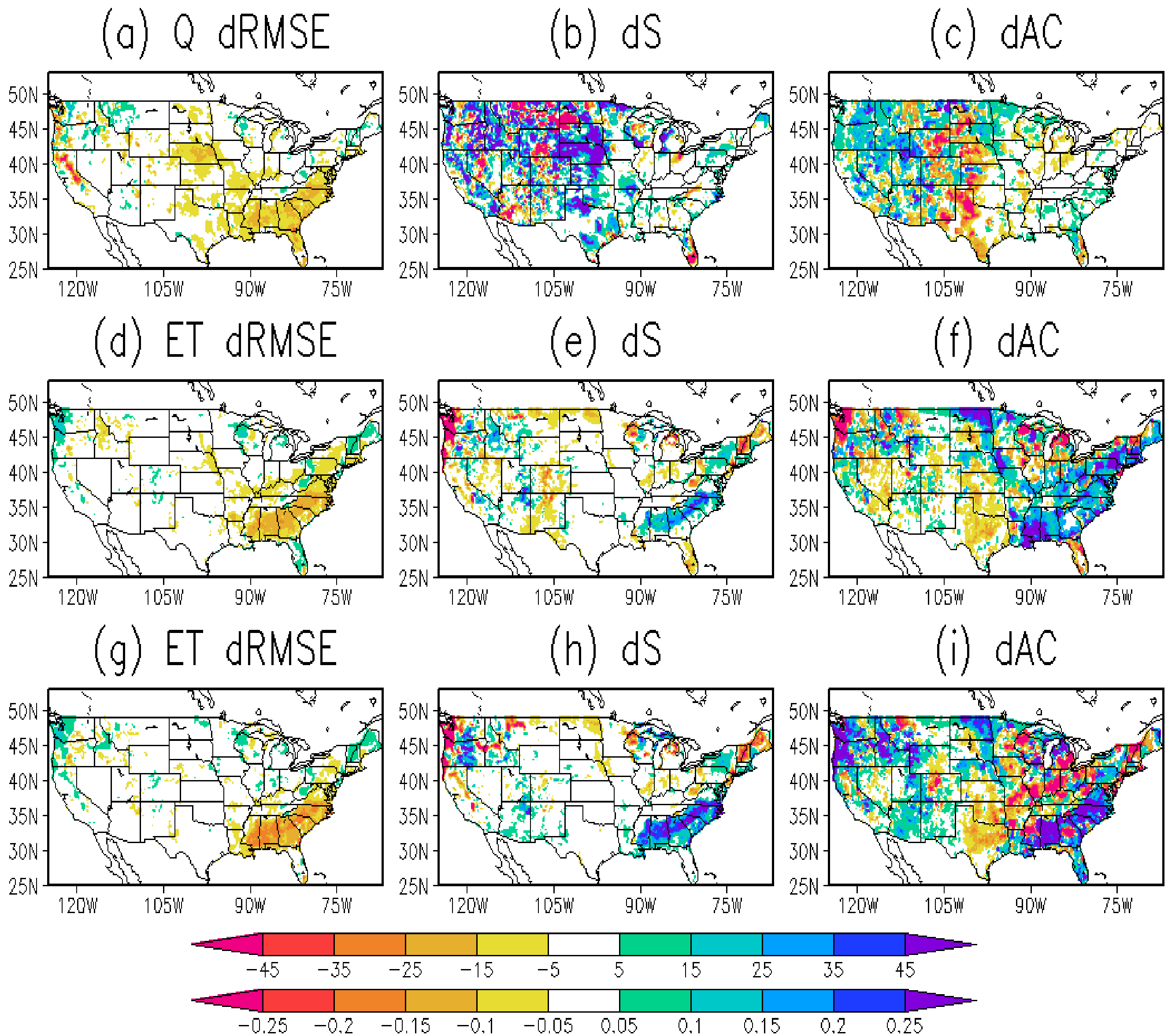


Figure 8: Statistics analysis for Q and ET for dRMSE, dS and dAC (dAC significance test at 95%)

For SNET and LNET AC, some improvement can be found in western mountains (due to snow cover and snow albedo effect on upward solar). For reason for dS (Lnet), see Figure S3. Mainly due to change canopy temperature calculation.

RMSE decrease and S increase in SE for SH and LH

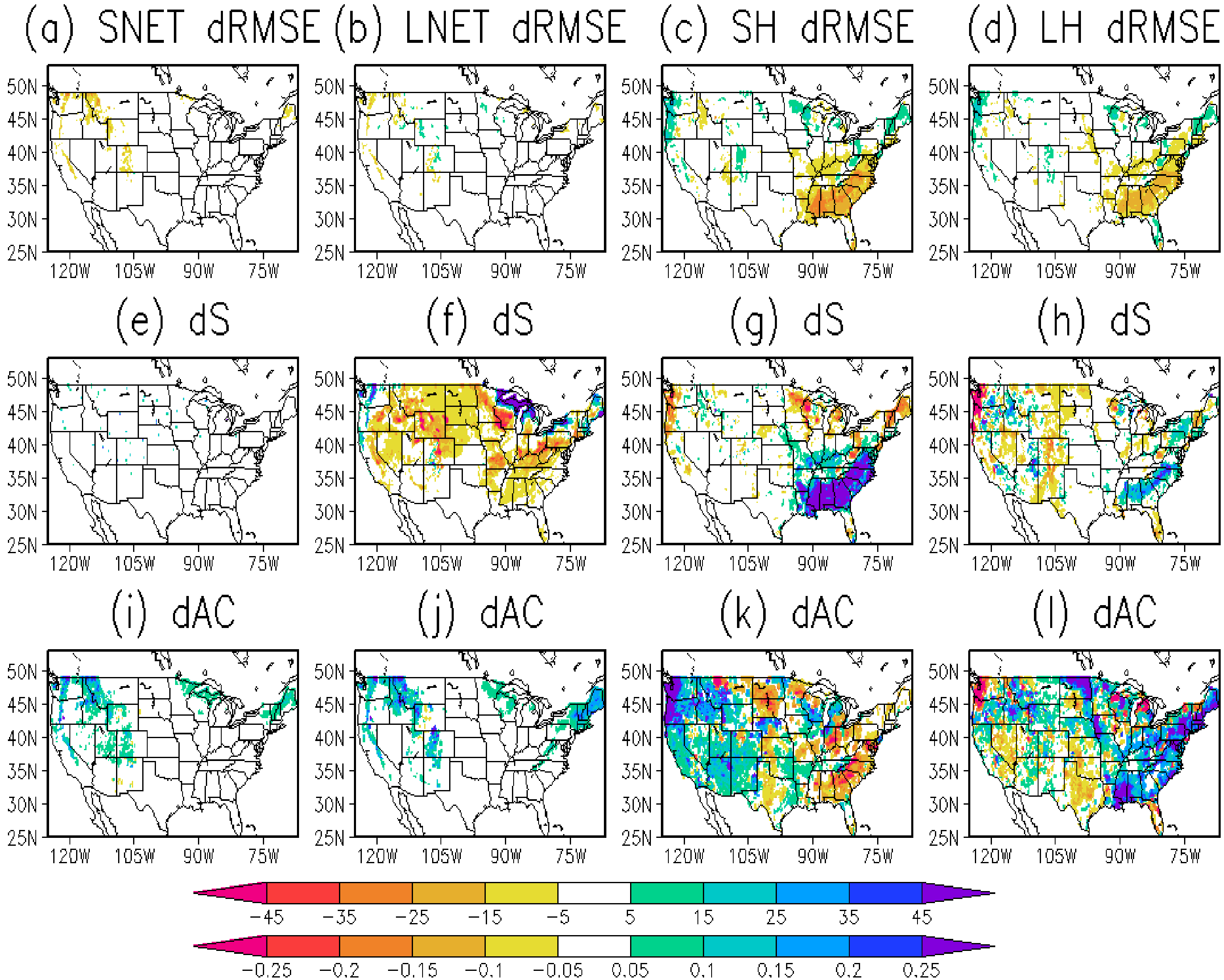


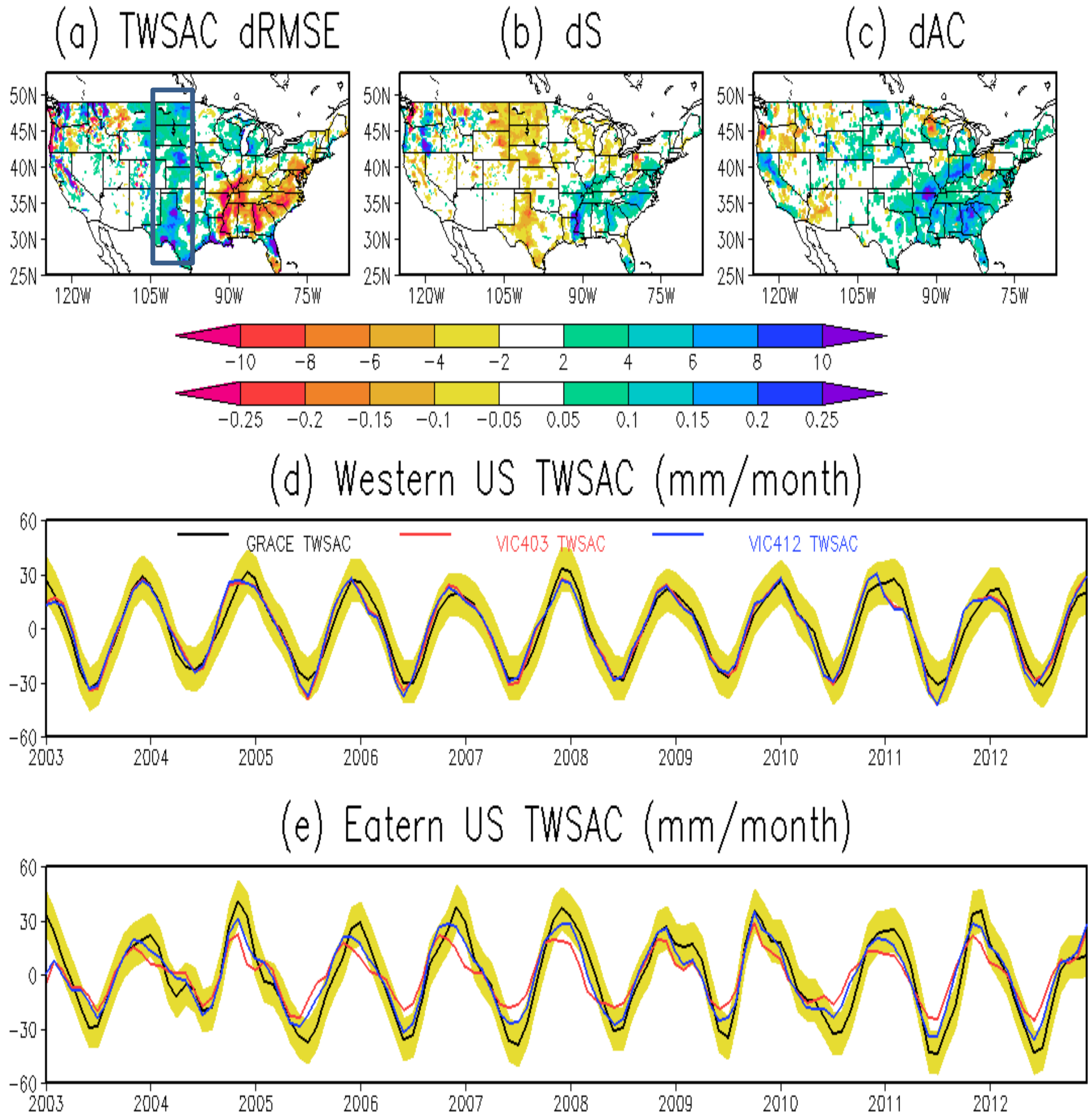
Figure 9: Statistics analysis for SNET, LNET, SH and LH (dAC significance test at 95%)

Figure 10: Statistics analysis for TWSAC

Overall VIC412 has significant improvement in SE using multiple metrics. However, S is largely reduced due to large RMSE increase.

This comes from both Q and ET treatment (see Figure S2, suggesting that a further investigation is needed.

Due to lack OBS in this region, either calibration process or model physics upgrade is still challenging the community



This suggesting that there are large uncertainties for our references due to different source data are used.

How to estimate these data errors and uncertainties is an important work not only for research but also application community.

Even so, using multiple observations and references is still needed in future

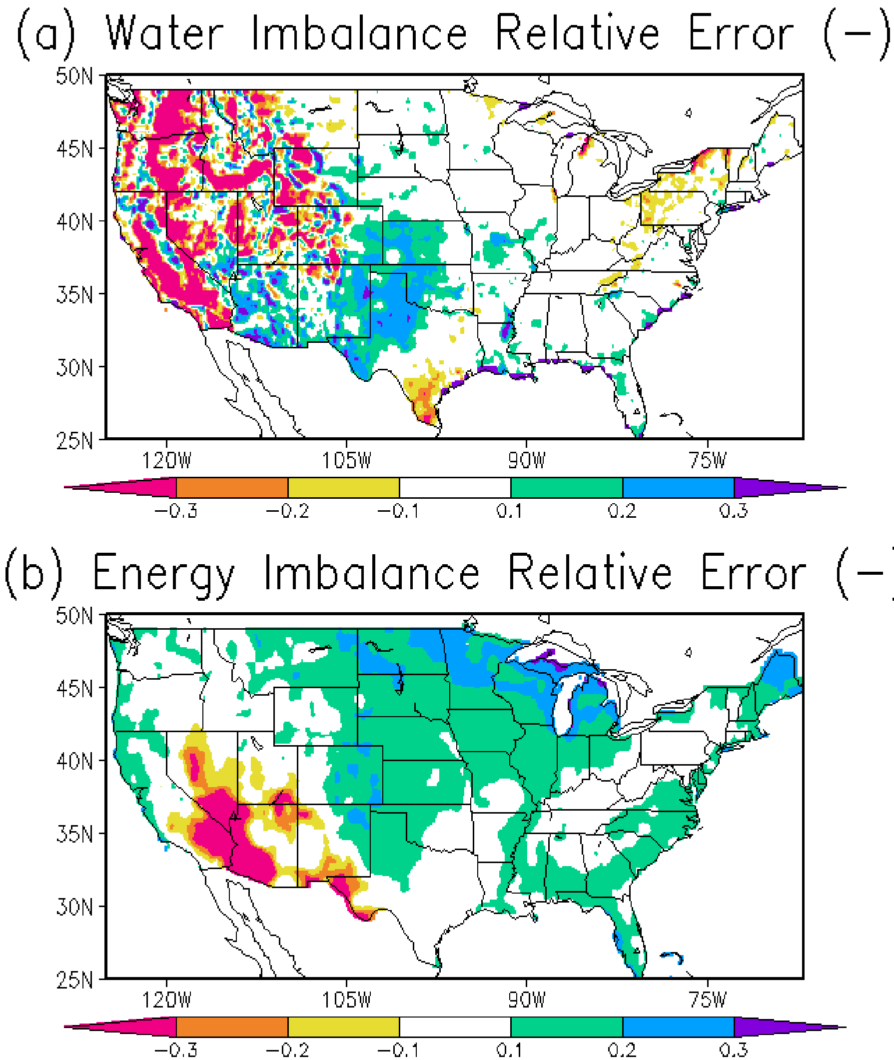


Figure 11: Water and energy imbalance issue due to various data sources

In general, model upgrade reduces RMSE, increases S and AC values for many regions except for Great Plains where there significant AC reduction (ET, SH, LH) and S reduction (Q).

This further suggests that this is a difficult region and needs community effort to enhance science understanding and observations

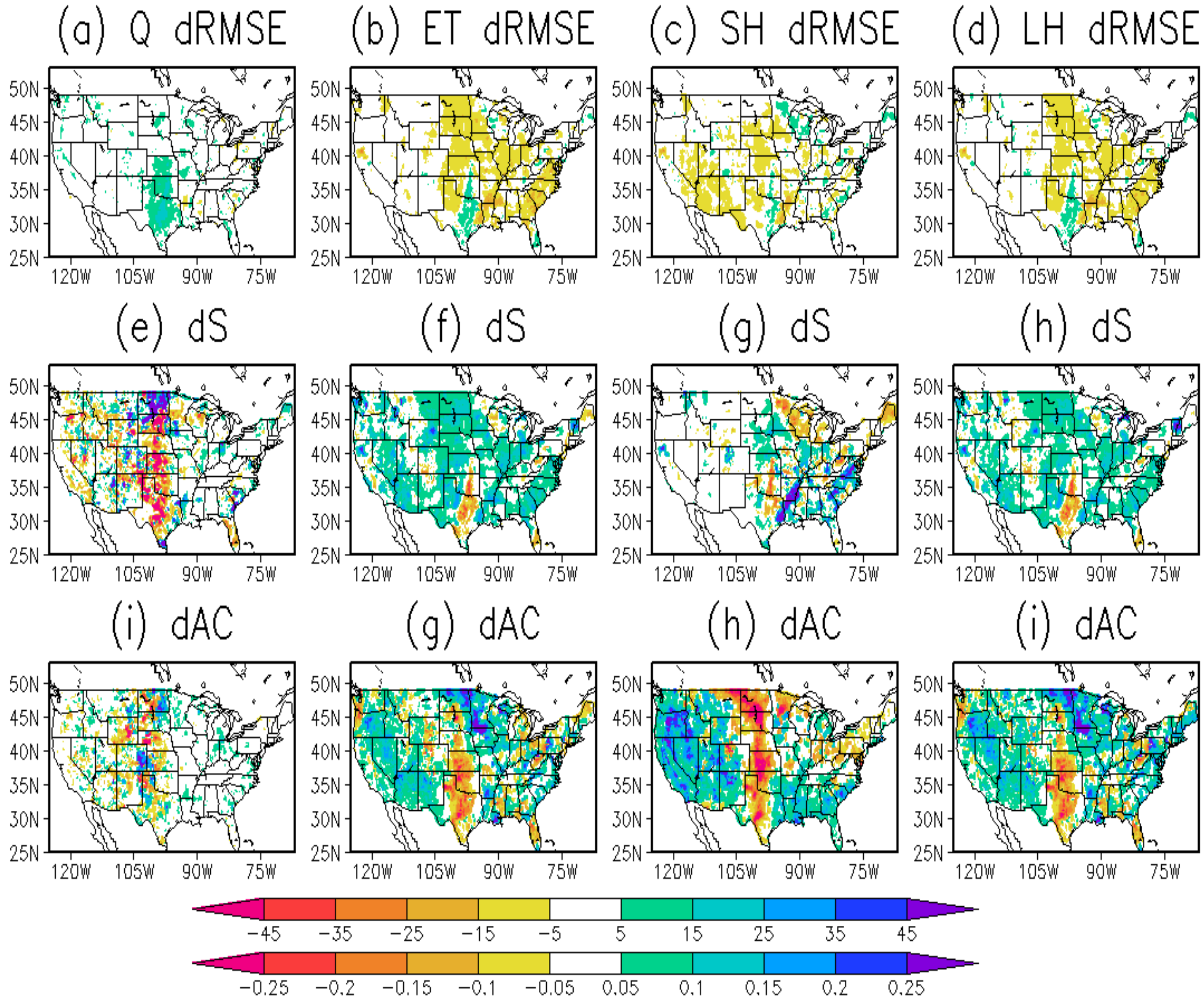


Figure 12: Improvement/deterioration due to model upgrade (VIC405 and VIC412 used the same and hydrology parameters)

Supplementary Materials

Also cite David Mocko's NLDAS science testbed results:

Top soil layer soil moisture AC (ARS, SCAN), and USGS streamflow (AC, NSE) for daily evaluation

Either put one Figure in manuscript or just cite his ppt file – need your comments here.

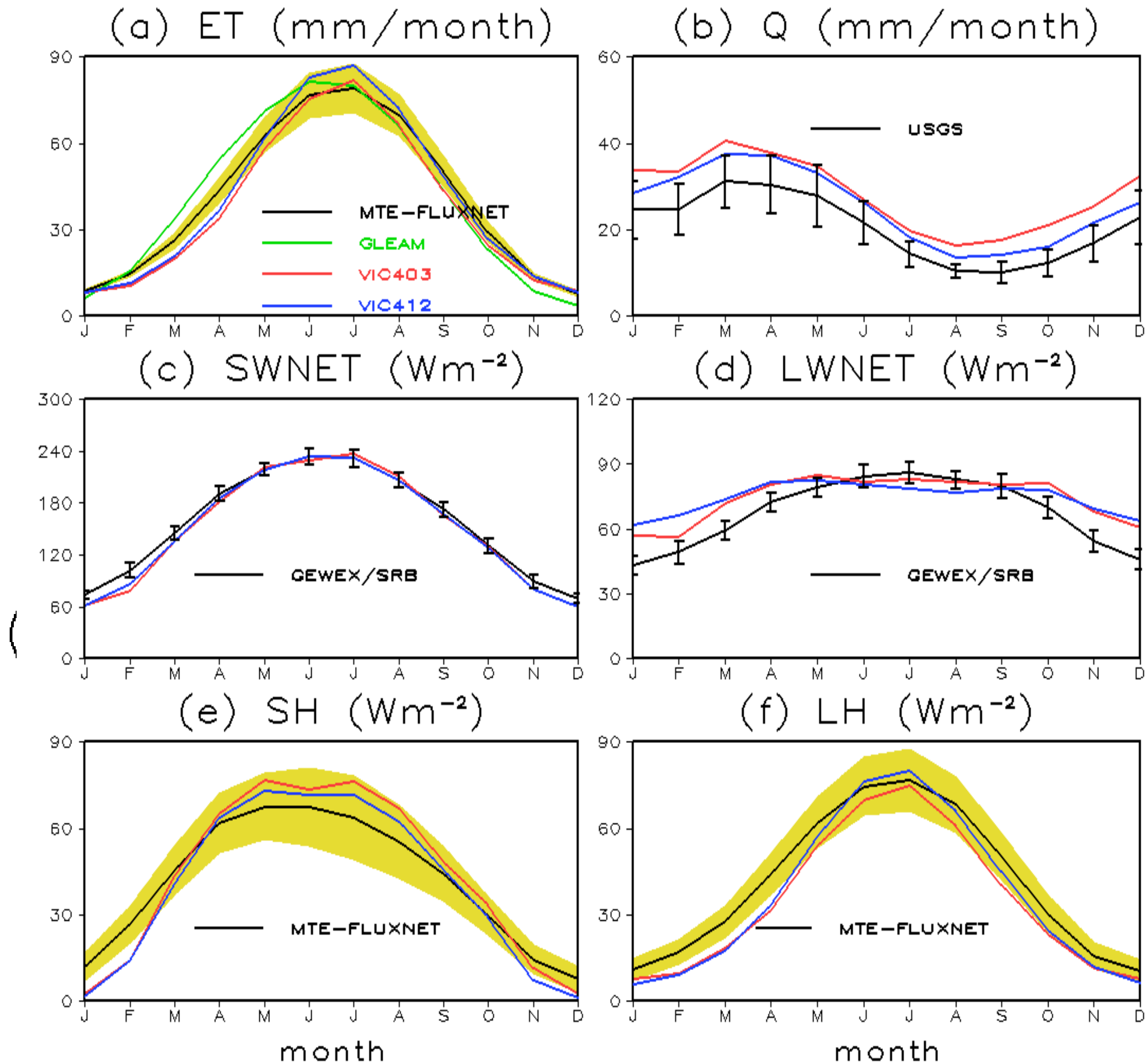


Figure S1: Mean seasonal cycles for (a) ET, (b) Q, (c) SWNET, (d) LWNET, (e) SH, and (f) LH over CONUS.

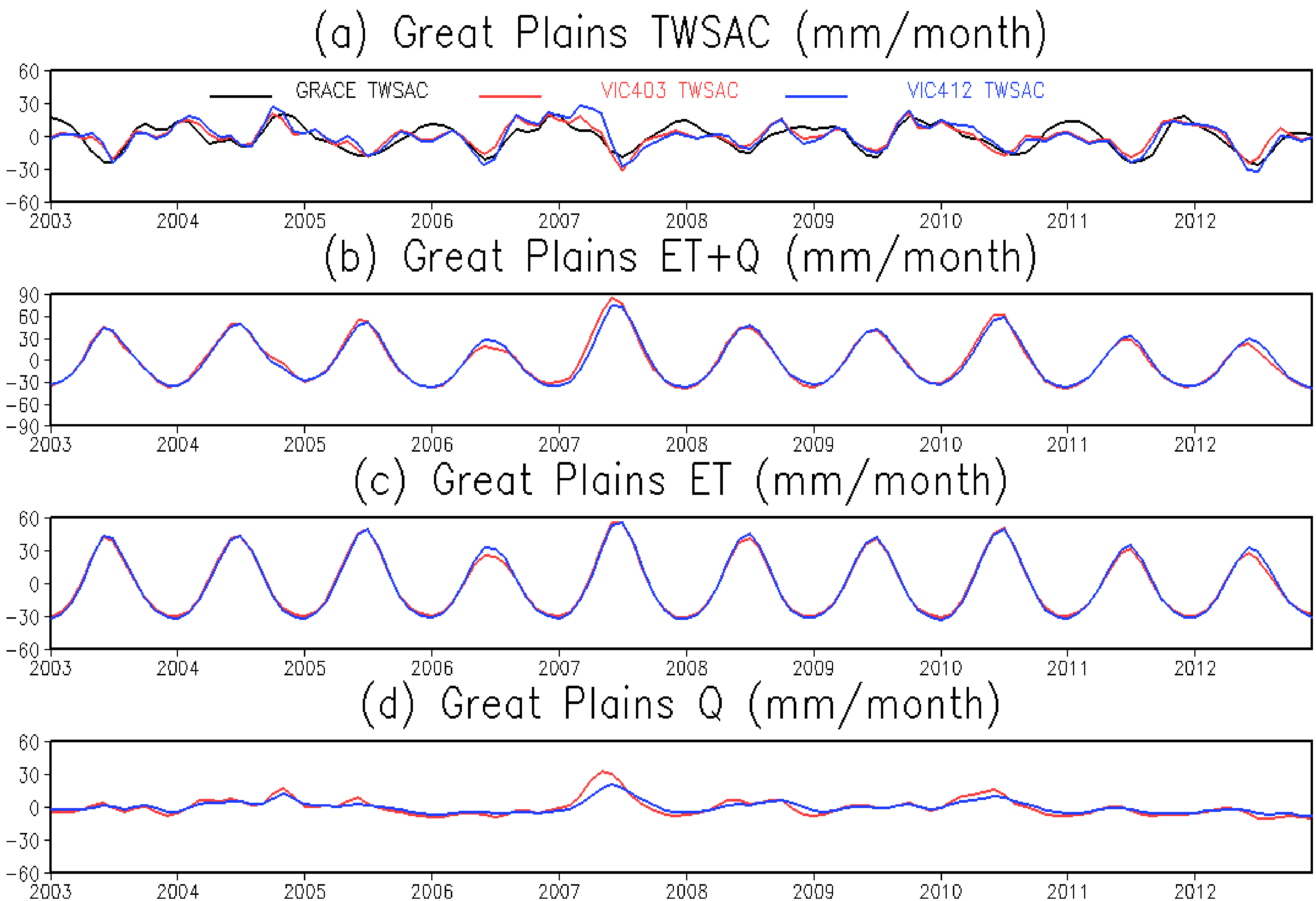
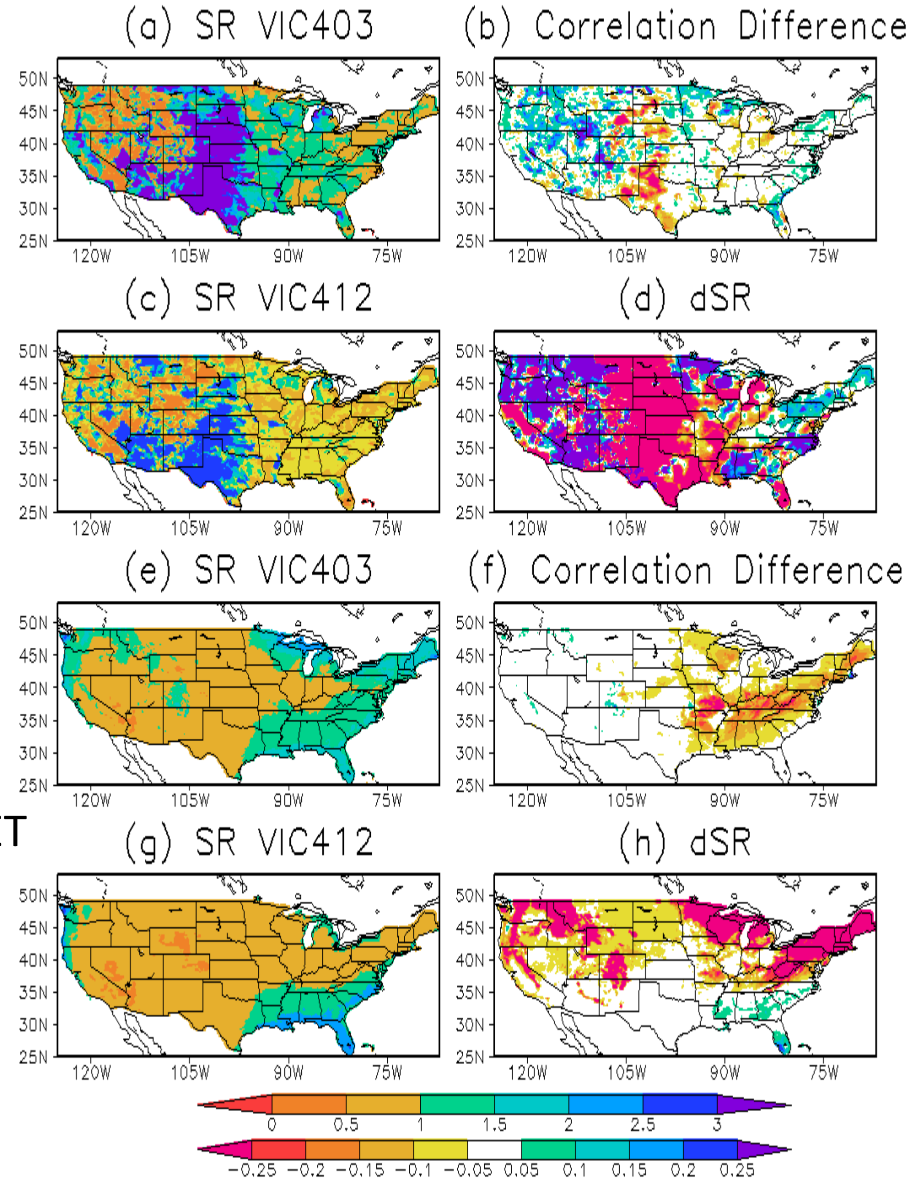


Figure S2: Area averaged anomalies over Great Plains (25-50^{o}N, 98-102^{o}W)

Q

LWNET



R increases, S always increases
SR depends if its value is larger than 1 or is smaller 1.

Perfect score is SR=1, and correlation=1.

SR>1, S improvement needs SR decrease;
SR<1, S improvement needs SR increase.

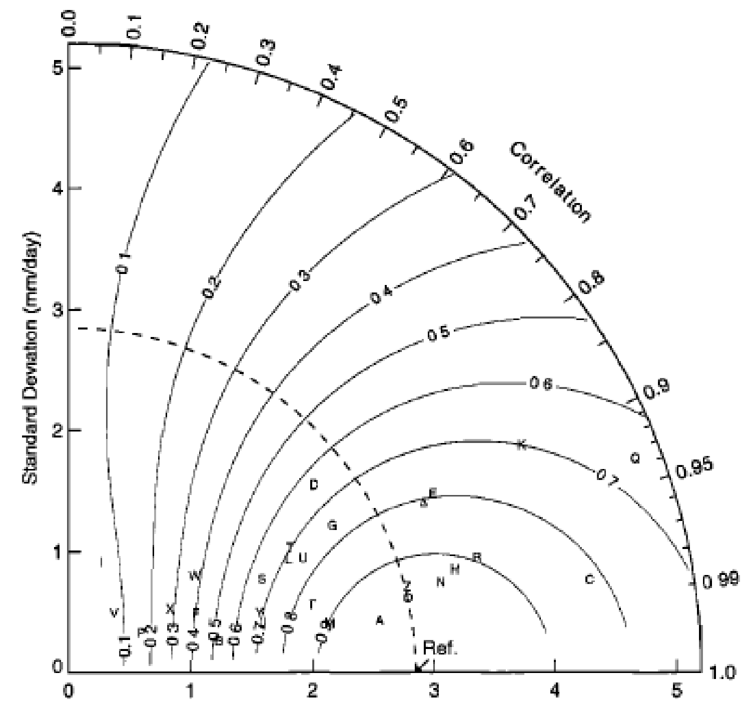


Figure 11. As in Figure 10, but for an alternative measure of skill defined by (5).

Figure S3: Separate analysis for Q (top four plots) and LWNET (bottom four plots) when standard deviation ratio, correlation, and dSR are used

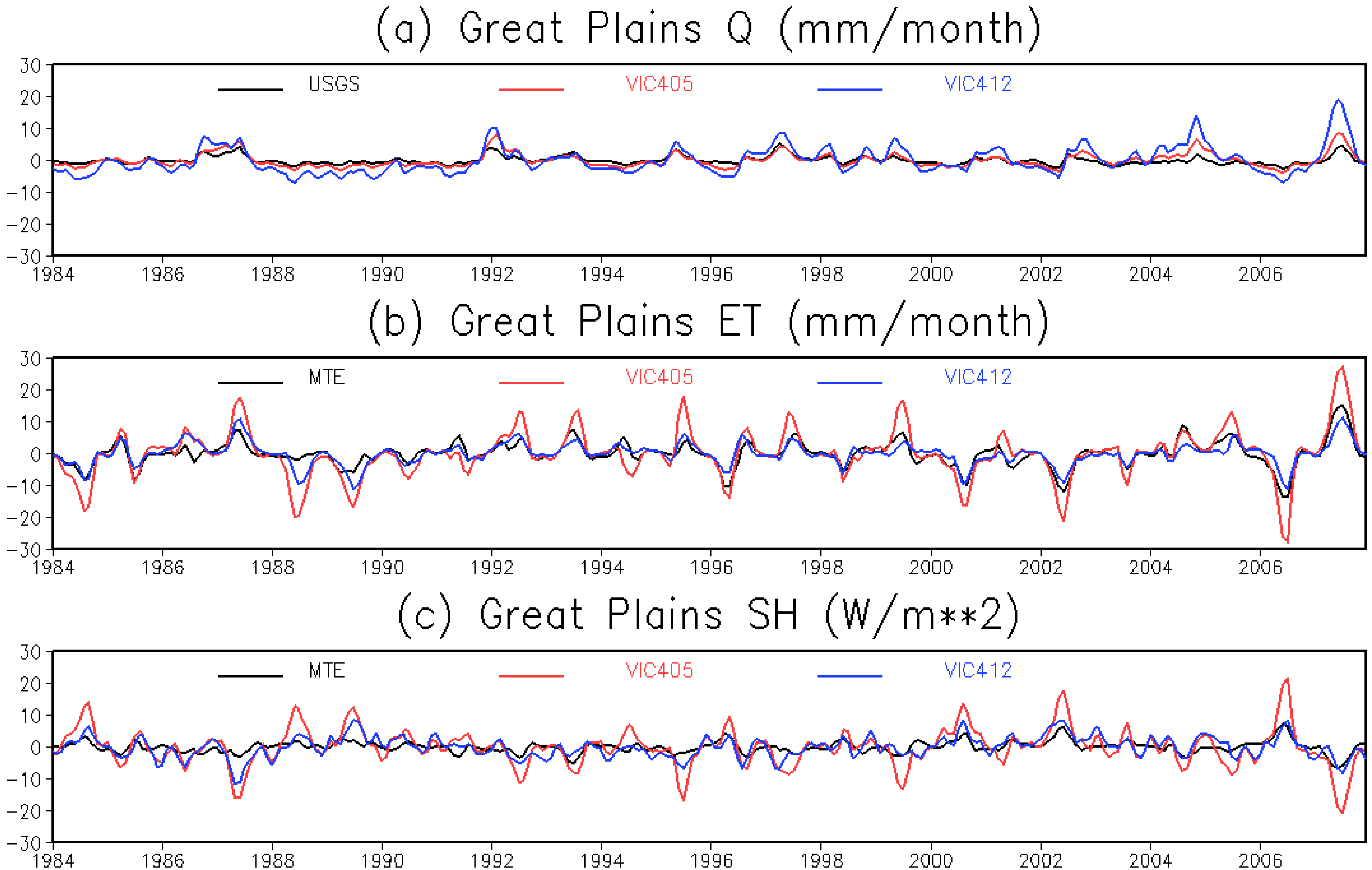


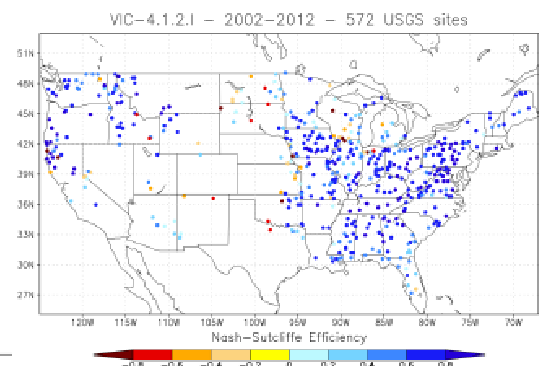
Figure S4: Area averaged anomalies over the Great Plains (25-50ºN, 98-102ºW) https://ldas.gsfc.nasa.gov/nldas/presentations/Mocko_NLDAS-Science-Testbed_2016-06-16.pdf

Daily Soil Moisture AC: 0.603 -> 0.492 at 4 ARS sites, 0.492->0.398 at 117 SCAN sites
Daily streamflow AC: 0.769->0.807 at 572 USGS small basins

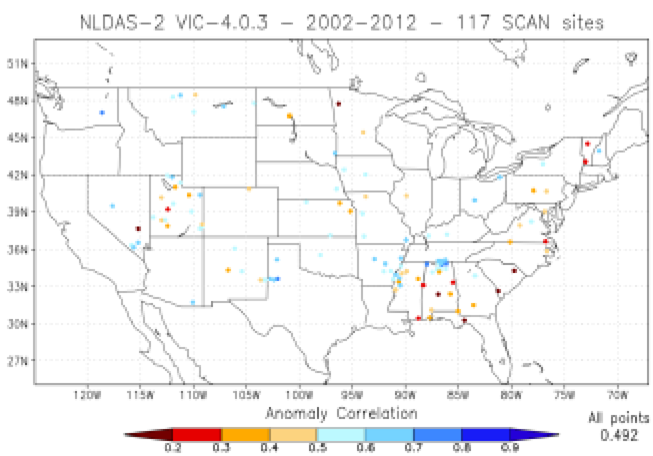
One Figure from David Mocko:

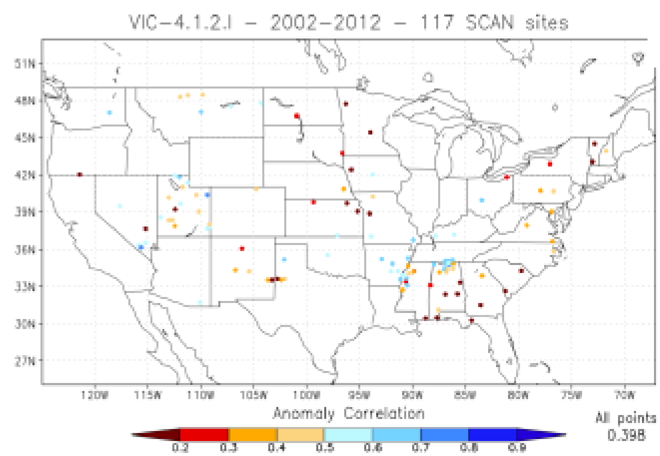
Daily streamflow evaluation

Daily NSE



(a) Spatial distribution of AC over 572 sites – VIC403, (b) VIC412-VIC403, (c) NSE VIC403, and (d) VIC412-VIC403 (-0.05-0.05 using grey color for difference)





I can put this Figure to supplementary materials

Conclusion

- Use VIC as an example to build a compressive evaluation framework for NLDAS system.
- Multivariate and multimeric evaluation evaluation is a more useful tool for NLDAS system and model community.
- As more and more in-situ measurements, remotely-sensed data, and reanalysis data become available, such a framework will become more important.
- This framework can be improved by adding multiple time scales and reference data uncertainty analysis into this framework