Benchmarking the next phase of the North American Land Data Assimilation System (NLDAS) using the Land Verification Toolkit (LVT)

David M. Mocko[1,2], Christa D. Peters-Lidard[1], Sujay V. Kumar[1,2], Shugong Wang[1,2], Kristi Arsenault[1,2], Grey Nearing[1,2], Youlong Xia[3,4], Michael B. Ek[3], and Jiarui Dong[3,4]

1 – Hydrological Sciences Laboratory, NASA/GSFC, Greenbelt, MD

2 – SAIC at NASA/GSFC, Greenbelt, MD, USA

3 – NOAA/NCEP/EMC, College Park, MD, USA

4 – IMSG at NCEP/EMC, College Park, MD, USA

This work is funded in part by the NOAA Climate Program Office's Modeling, Analysis, Predictions, and Projections (MAPP) Program

Acknowledgements:

Gabrielle De Lannoy, Ben Livneh, Yudong Tian, Rolf Reichle, Qing Liu, Augusto Getirana, Yuqiong Liu, Sarith Mahanama, Mike Cosh

Presentation Outline

- Introduction to the North American Land Data Assimilation System (NLDAS) project and the new/upgraded LSMs for the next phase on NLDAS
  - **Noah-3.3, Catchment/Fortuna-2.5 (shown here); SAC-HTET-3.5.6/SNOW-17, VIC-4.1.2.1 (in LIS-7)**
- Simulations with the Land Information System (LIS)
  - **Introduction to the LIS software framework**
- Model evaluations and benchmarks using the Land Verification Toolkit (LVT)
  - **Soil moisture, Surface fluxes, Snow, Streamflow evaluations using LVT**
  - **Regression model development and evaluation**

NLDAS Land Surface Models (LSMs)

- NLDAS Phase 2 is currently running routinely in near-real time (~3.5-day lag) to drive a suite of LSMs from the meteorological (**Noah-2.8 and Mosaic**) and hydrological (**Sacramento [SAC/SNOW-17] and VIC-4.0.3**) communities.
- The NLDAS-2 LSMs have been extensively evaluated in several papers by Xia et al. for soil moisture/temps, streamflow, fluxes, etc.
- For the next phase of NLDAS, new and upgraded LSMs are being run using the NASA-developed LIS software framework, including the use of data assimilation capabilities within LIS.
- All LSMs are run on a 1/8th deg. resolution CONUS domain, including parts of Canada/Mexico (25-53° N; 125-67° W).
- A 60-year spin-up of the soil states was performed, followed by 34-year simulations from Jan 1979 – Dec 2012.

Reference(s): Xia et al. (2012a&b) – NLDAS-2 introduction and streamflow evaluation – *JGRa*; Xia et al. (2013) – soil temperature evaluation – *J. Appl. Meteor. Climatol.*; and numerous

NLDAS Land Surface Models (LSMs)

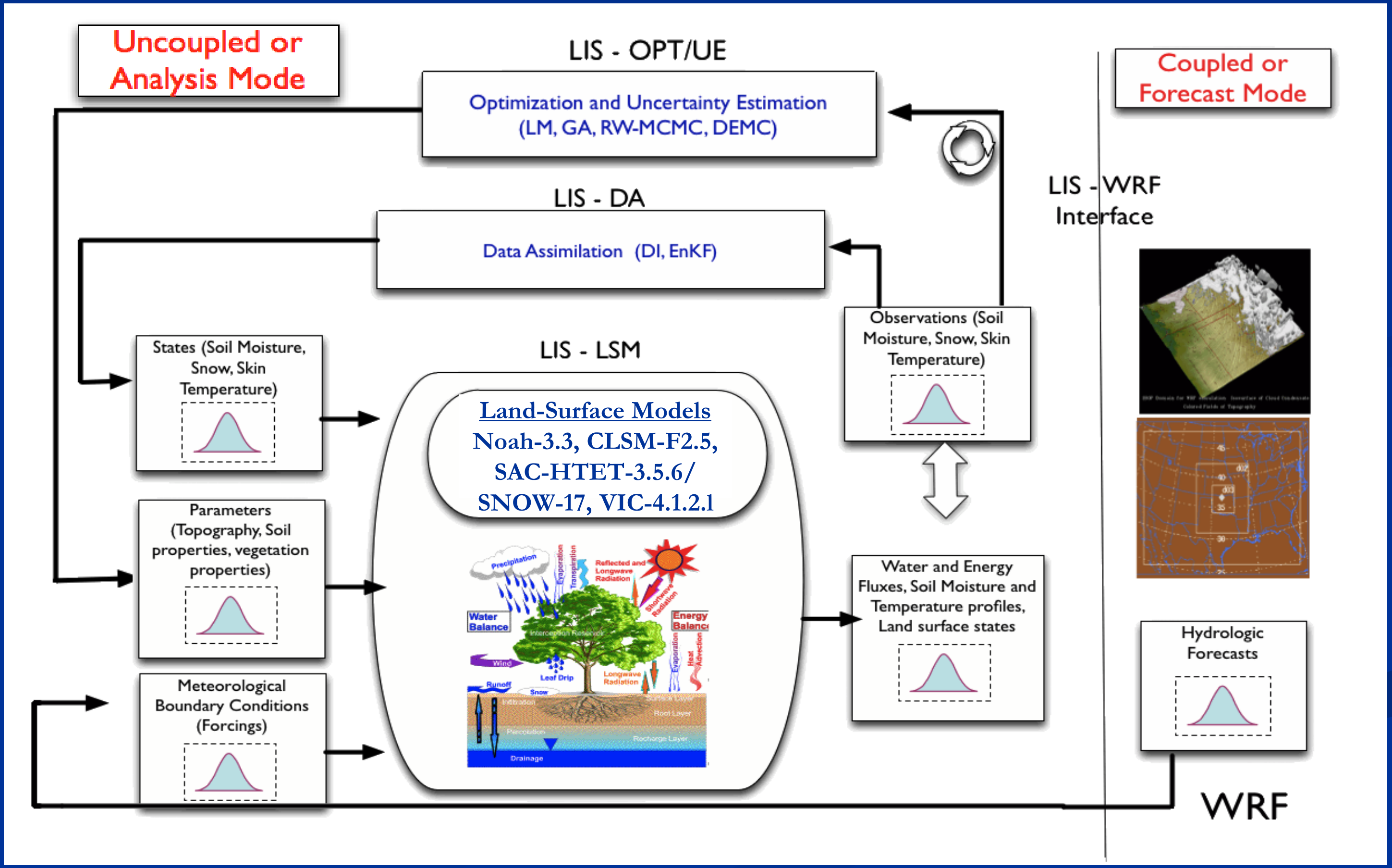
NLDAS-2	Major LSM changes	Next phase of NLDAS	References
Noah-2.8	• Common code by NCAR/NCEP • Warm season updates • Snow physics upgrade	Noah-3.3	Chen et al. (1996, JGR); Ek et al. (2003, JGR); Wei et al., 2012, HP); Livneh et al., 2010, J. Hydromet.)
Mosaic	• Topographic catchments instead of 1-D soil moisture layers • 3 soil moisture regions: saturated, sub-saturated, and wilting	Catchment/ Fortuna-2.5 (CLSM-F2.5)	Koster et al. (2000, JGR); Reichle et al. (2011, J. Climate); same version of code as for MERRA-Land
VIC-4.0.3	• Canopy energy balance • Snowpack improvements	VIC-4.1.2.1	Liang et al. (1994, JGR); Gao et al. (2010, book chapter)
SAC/ SNOW-17	• Distinct soil layers for soil moisture/temps (HT) • Includes the Noah LSM's evapotranspiration physics (ET)	SAC-HTET-3.5.6/ SNOW-17	Burnash et al., (1973); Anderson (1973); Koren et al. (2007, 2010, NOAA Tech Memos)

The Land Information System (LIS)

- LIS is a flexible land-surface modeling and data assimilation framework developed with the goal of integrating satellite- and ground-based observed data products with land-surface models.

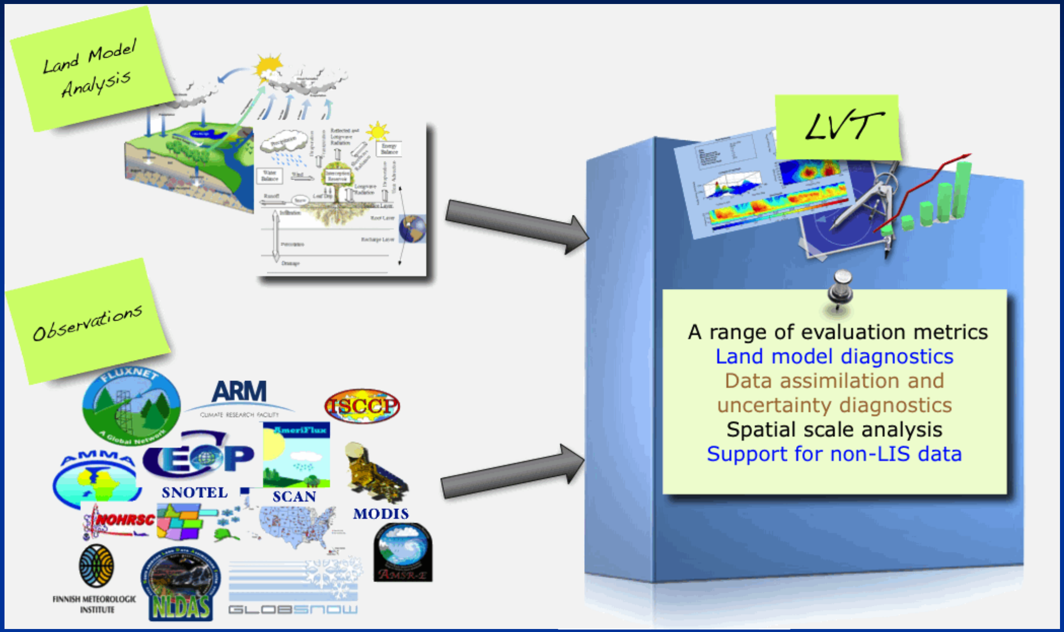
Data Assimilation of: Soil Moisture, SWE, SCF, TWS

1979-present NLDAS-2 Forcing and Parameters



Reference(s): Kumar et al. (2006) in *Environmental Modelling & Software*
Peters-Lidard et al. (2007) in *Innovations in Systems and Software Engineering*

The Land Verification Toolkit (LVT)



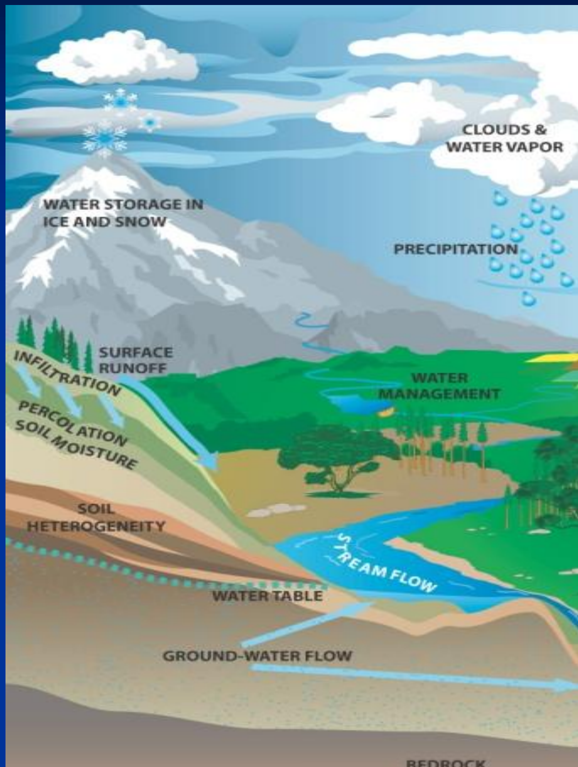
Metric Class	Examples
Accuracy	RMSE, Bias, Correlation
Ensemble	Mean, Standard deviation, Likelihood
Uncertainty	Uncertainty importance
Information theory	Entropy, Complexity
Data assimilation	Mean, variance, lag correlations of innovation distributions
Spatial similarity	Hausdorff distance
Scale decomposition	Discrete wavelet transforms

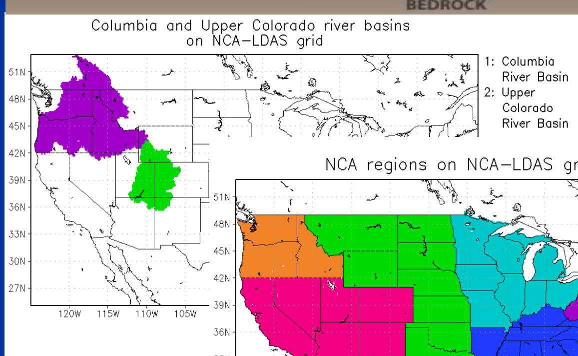
- LVT is a NASA-developed open-source software framework developed to provide an automated, consolidated environment for systematic land surface model evaluation and benchmarking
- Includes support for a range of in-situ, remote-sensing, and other model and reanalysis products in their native formats

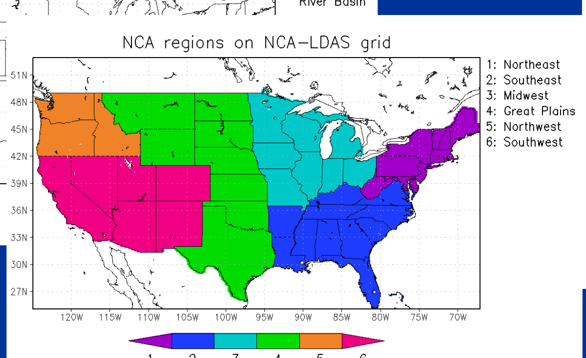
Reference: Kumar et al. (2012): Land surface Verification Toolkit (LVT) – A generalized framework for land surface model evaluation. *Geosci. Model. Dev.*

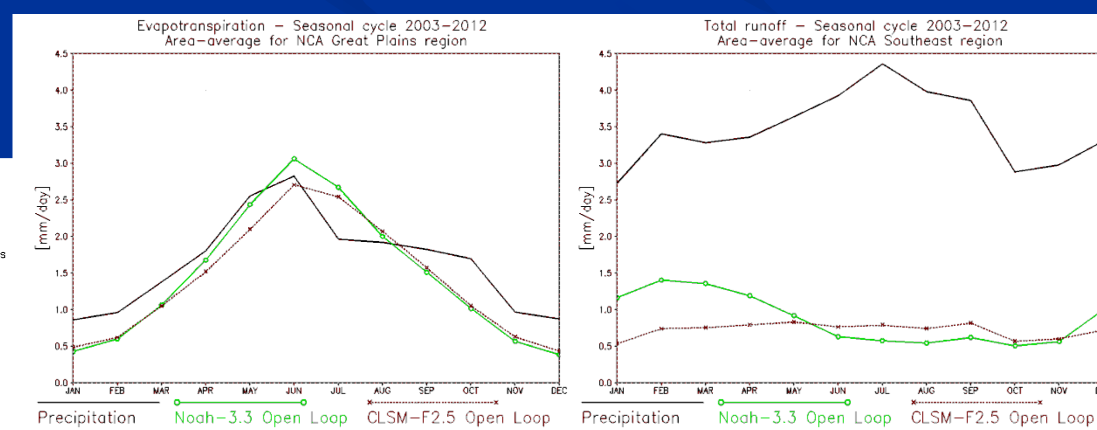
Benchmarking philosophy

1) Compare against available independent data – evaluating against every component of the water cycle, including budgets, balances, and combination variables (SWE/P, E/P, Q/P).
2) Investigate if the LSMs provide additional information compared to a regression model (following on the GLASS PALS Land Surface Model Evaluation Benchmarking – PLUMBER – Project).

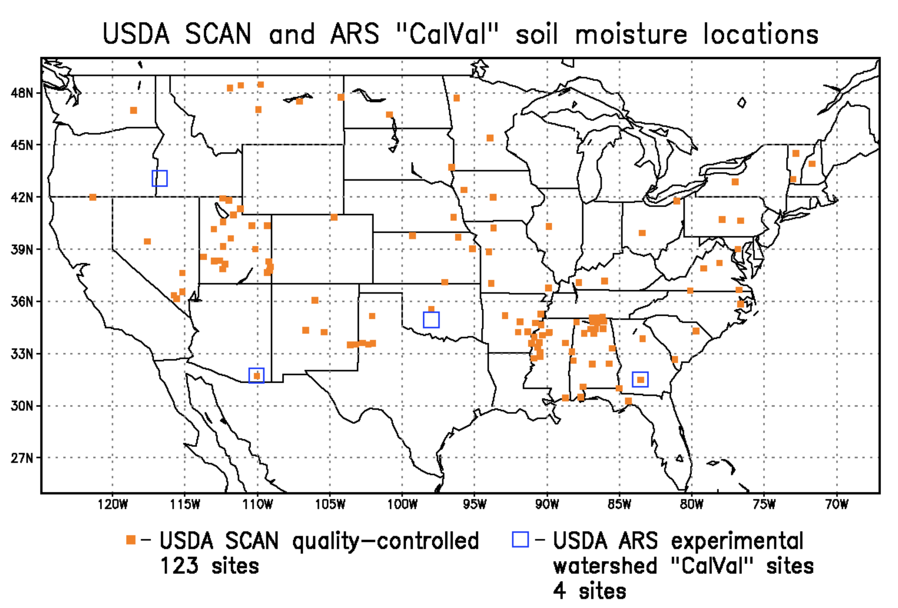








Evaluation of soil moisture fields



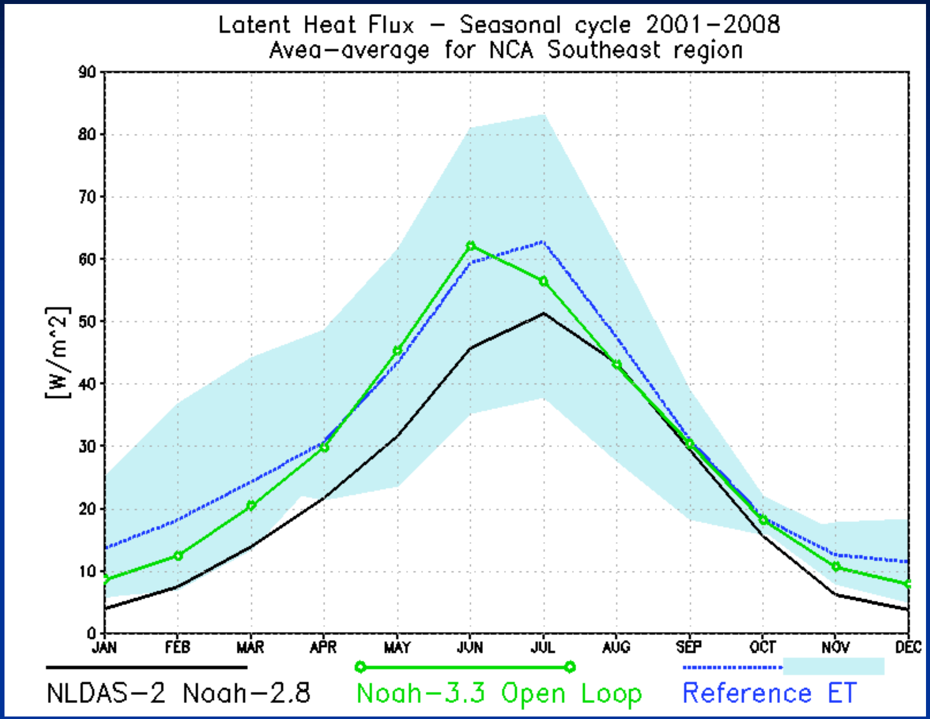
LSM (version)	ARS	SCAN
NLDAS-2 Mosaic	0.758	0.661 +/- 0.025
NLDAS-2 SAC	0.757	0.701 +/- 0.027
NLDAS-2 VIC-4.0.3	0.592	0.495 +/- 0.029
NLDAS-2 Noah-2.8	0.694	0.634 +/- 0.028
Noah-3.3 OL	0.711	0.614 +/- 0.024
CLSM-F2.5 OL	0.606	0.432 +/- 0.026

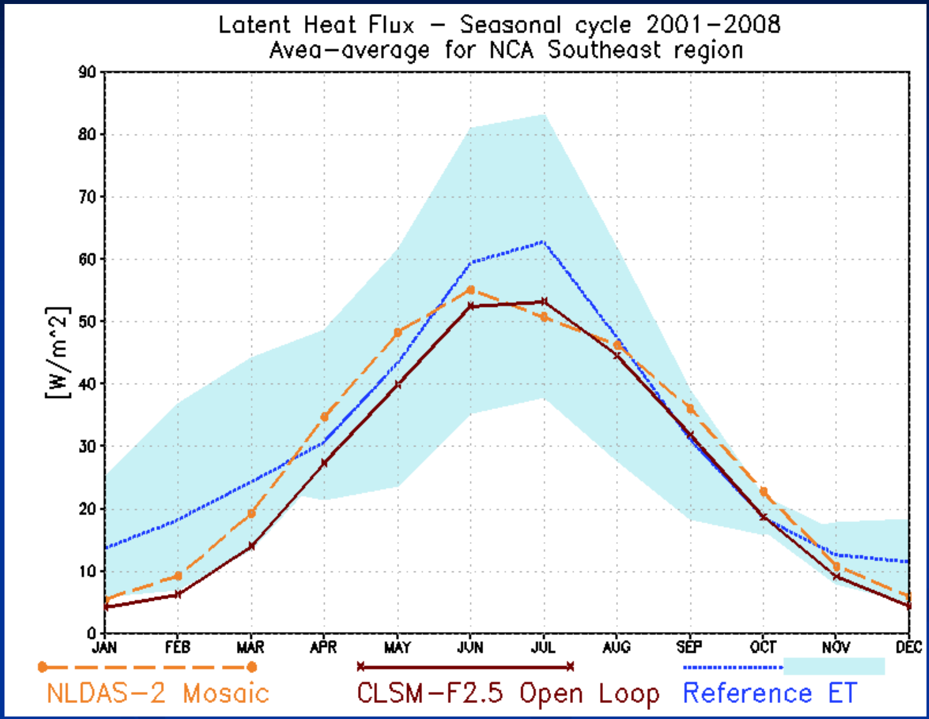
Anomaly R values

The surface soil moisture was evaluated against quality-controlled soil moisture observations at 5-cm depth from the USDA SCAN (123 sites) 2000-2012 and ARS "CalVal" networks (4 sites) for 2001-2011. Noah-3.3 OL (Open Loop) shows an improvement over NLDAS-2's Noah-2.8 for the ARS sites, while showing a degradation for the SCAN sites. This NLDAS configuration of the CLSM-F2.5 OL does not do as well as NLDAS-2 Mosaic, but has improvements to other aspects of the water balance (shown later). Data assimilation of remotely-sensed soil moisture and snow for Noah-3.3 and of GRACE TWS anomalies for CLSM-F2.5 has shown improvements to the soil moisture evaluations (shown earlier this week by Kumar et al. and by Kumar et al. (2014, *JHM*, Special Collection on "Advancing Drought Monitoring and Prediction")).

Reference(s): Jackson et al. (2012) – ARS – *IEEE TRGS*; Schaefer et al. – SCAN – *J. Atmos. Oceanic Technol.*; Liu et al. (2011) – soil moisture skill with land DA – *JHM*

Evaluation of gridded surface fluxes

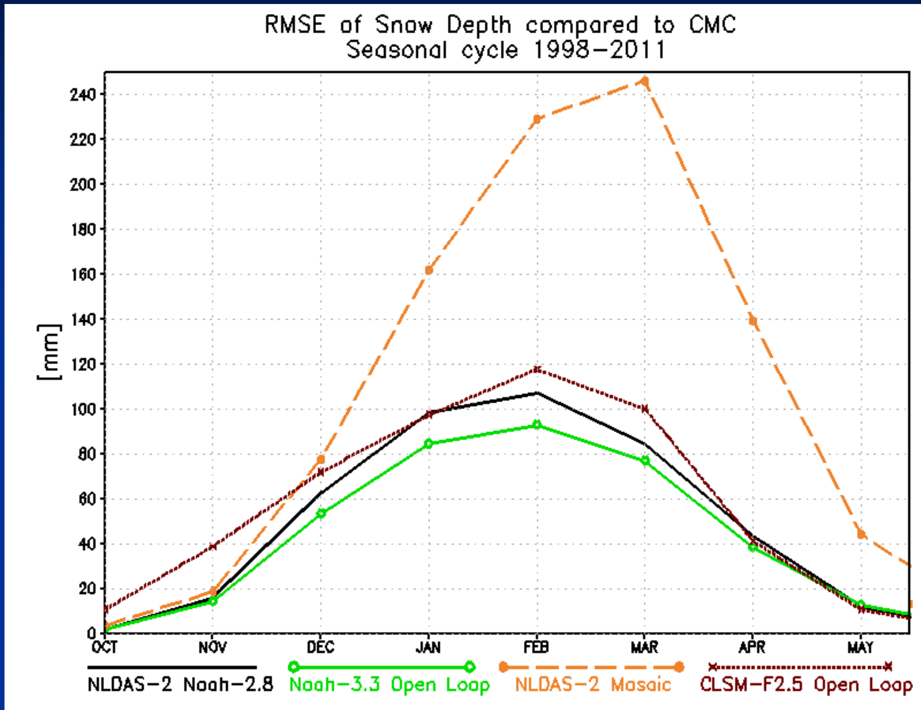


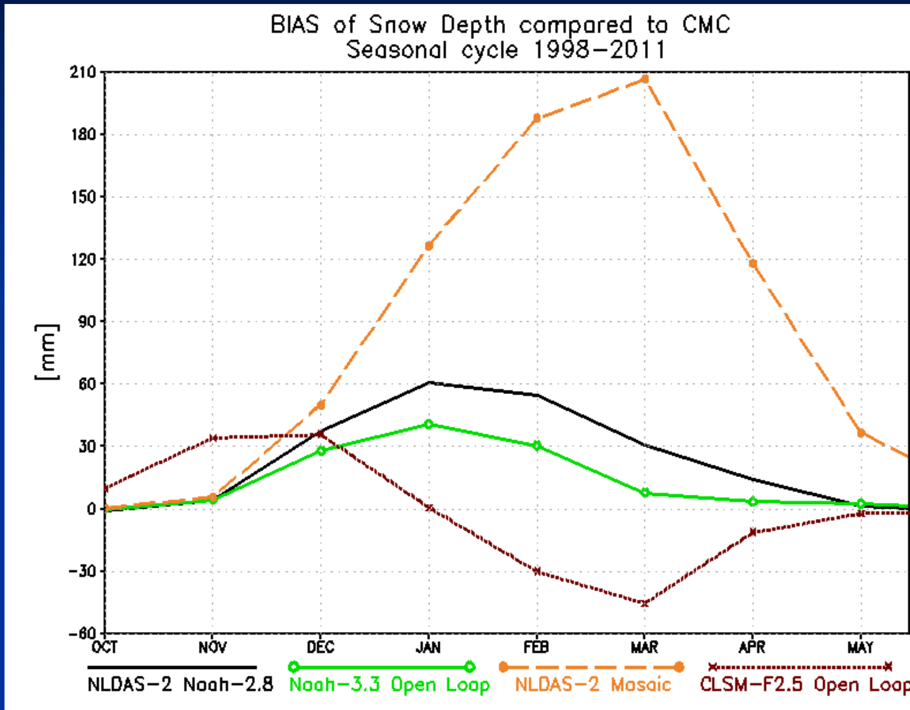


Gridded monthly surface fluxes based on FLUXNET surface observations, on MODIS retrievals (both MOD16 and UW ET), and on thermal remote sensing (ALEXI) are used to evaluate the fluxes from the NLDAS-2 LSMs and the new simulations. Shown is the seasonal cycle for the **NCA Southeast** region for 2001-2008, with the **mean**/**range** of the four gridded reference products shown in blue. The left panel shows the Noah-3.3 OL with higher ET in the spring and early summer compared to NLDAS-2 Noah-2.8. The right panel shows CLSM-F2.5 OL similar to NLDAS-2 Mosaic. See Peters-Lidard et al. (2011, *HP*) for more detailed ET evaluations.

Reference(s): Jung et al. (2009) – FLUXNET – *Biogeosci.*; Mu et al. (2011) – MOD16 – *Rem. Sens. Environ.*; Tang et al. (2009) – UW ET – *JGR*; Anderson et al. (2007) – ALEXI – *JGRa*

Evaluation of snow depth





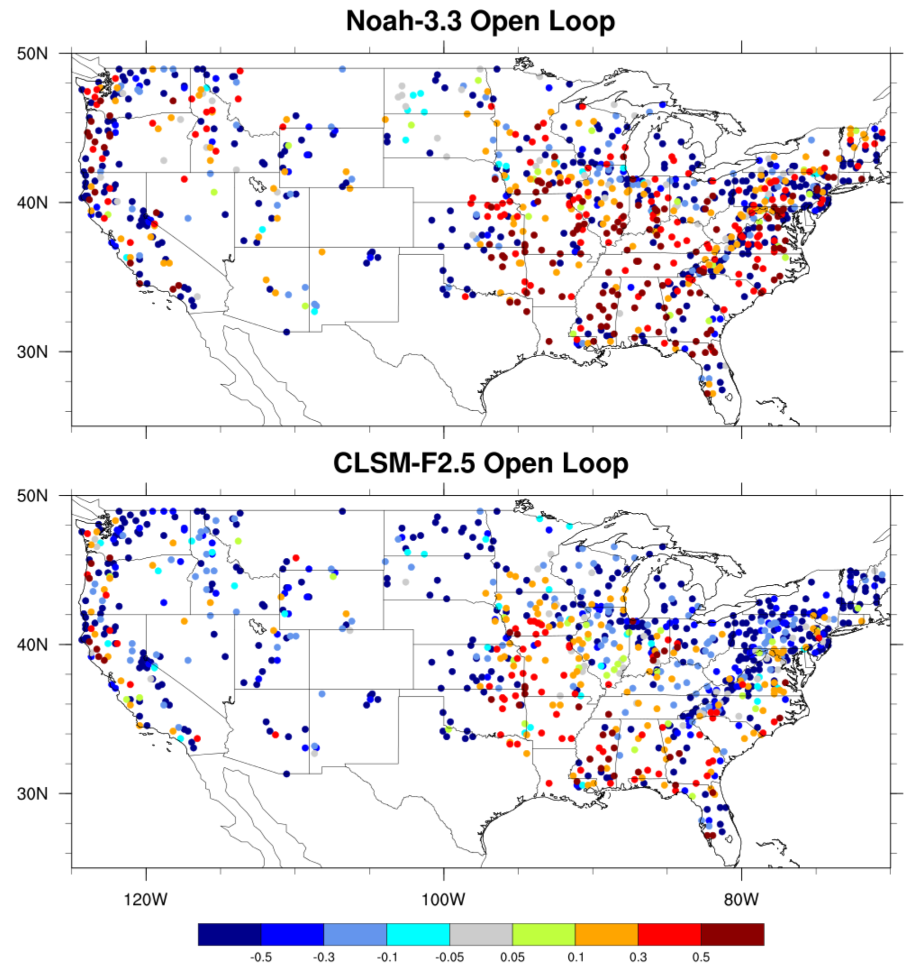
Gridded snow depth observations/analyses are used to evaluate the simulated snow depths in NLDAS. The reference products are Canadian Meteorological Centre's (CMC) daily snow depth analysis, the NWS's NOHRSC SNOw Data Assimilation System (SNODAS), and the Global Historical Climatology Network (GHCN). These figures (using CMC as the reference dataset for 1998-2011) show reductions in bias (left) & RMSE (right) from NLDAS-2 Noah-2.8 to Noah-3.3. CLSM-F2.5 also has a significantly lower bias and RMSE when compared to NLDAS-2's Mosaic. DA of SWE and SCA (Kumar et al.) has also been shown to further reduce the bias and RMSE.

Reference(s): Menne et al. (2012) – GHCN – *J. Atmos. Oceanic Technol.*; Brown and Brasnett (2010) – CMC – *Environ. Canada*; Barrett (2009) – SNODAS – *NSIDC, Boulder, CO*

Evaluation of streamflow

LSM (version)	AC	RMSE (m3/s)
NLDAS-2 Mosaic	0.672	57.8 +/- 16.9
NLDAS-2 SAC	0.763	63.7 +/- 21.1
NLDAS-2 VIC-4.0.3	0.758	106.1 +/- 40.1
NLDAS-2 Noah-2.8	0.715	118.0 +/- 42.9
Noah-3.3 OL	0.678	47.4 +/- 13.5
CLSM-F2.5 OL	0.574	44.2 +/- 13.8

(Left) Anomaly correlation (AC) and RMSE in (m3/s) for streamflow (from routed runoff) for 2000-2012 against USGS small unregulated basins. Noah-3.3 OL and CLSM-F2.5 OL have lower RMSE, but also slightly lower AC values as compared to the NLDAS-2 LSMs. Again, DA of soil moisture, snow, and terrestrial water storage has been shown to improve the evaluations of streamflow. We have also used a large-basin naturalized streamflow datasets for evaluations (see Mahanama et al., 2012).



(Right) Nash-Sutcliffe Efficiency against 939 USGS small unregulated basins, 2000-2012.

Reference(s): Lohmann et al. (2004) – NLDAS router and USGS streamflow obs – *JGR*; Getirana et al. (2012) – HyMAP router – *JHM*; Mahanama et al. (2012) – Naturalized streamflow – *JHM*;

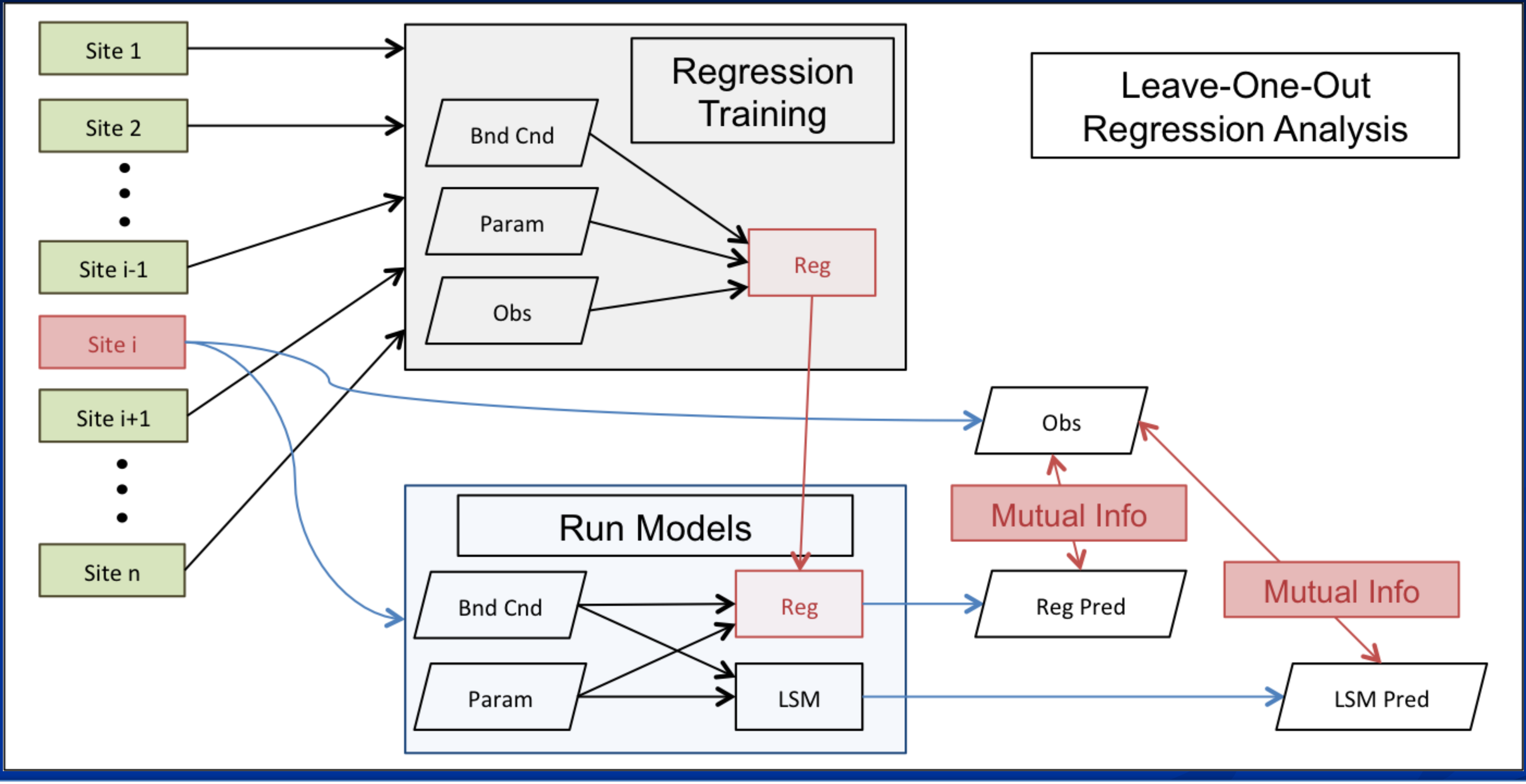
Model inputs vs. Model physics

Information from: **Model Inputs** vs. **Model Physics**

"for the most part, the models *under-utilise* the information available to them" - (Abramowitz, 2005; GRL)

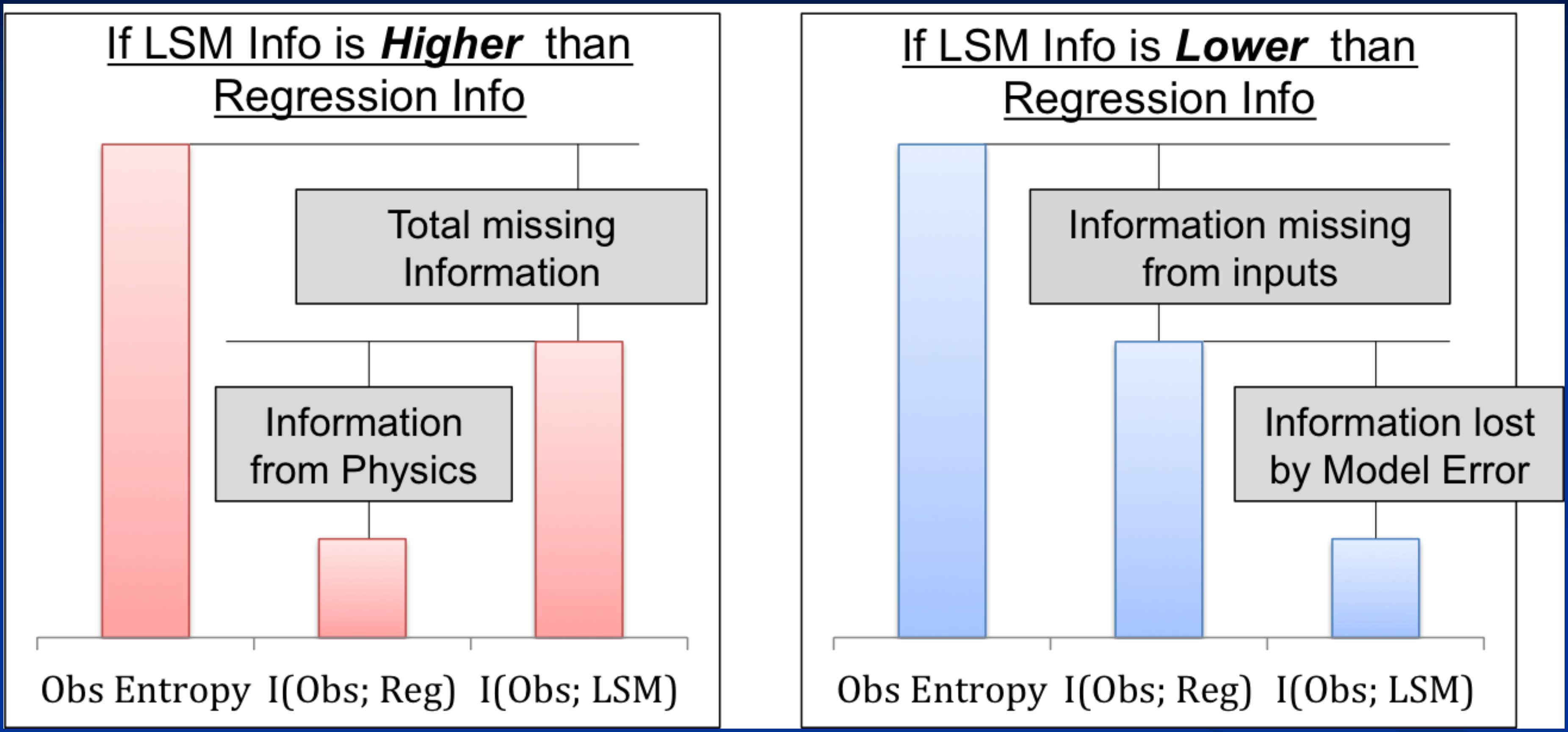
- How to measure information use efficiency?
  - Abramowitz (2005; 2012) showed that regressions on the boundary conditions often out-perform physics-based LSMs. **Implies that models do not use all available information.**
  - Gong et al. (2013) measured information *lost by the model*.
  - Nearing et al. (in review) measured information *provided by model physics*.
- Use a Leave One Out (LOO) regression model

Regression analysis



With observations at "n" Sites, Site "i" is left out of the training of the regression model, which uses the inputs and observations at each other site. After training, the inputs and observations at Site "i" are used to determine information obtained from the observations, the regression model, and the physics of the land-surface model(s). This is repeated for all "n" Sites.

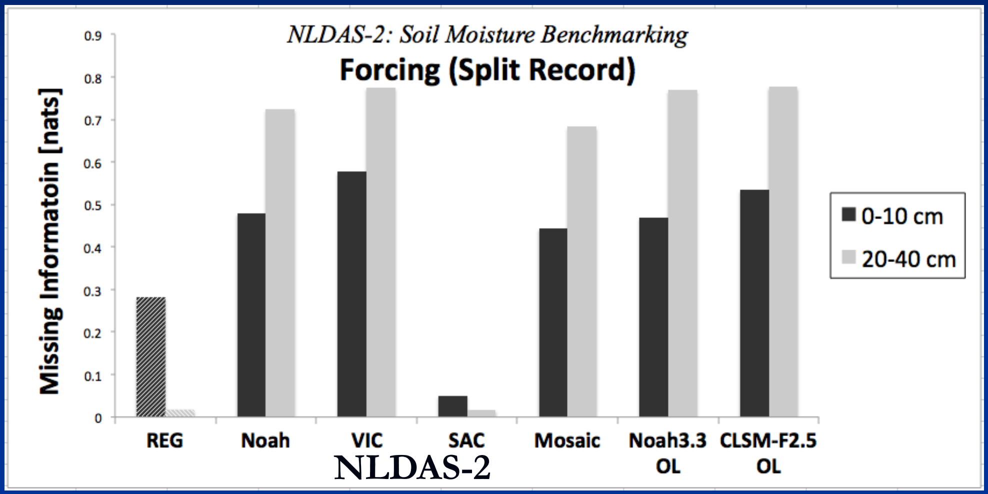
Regression analysis

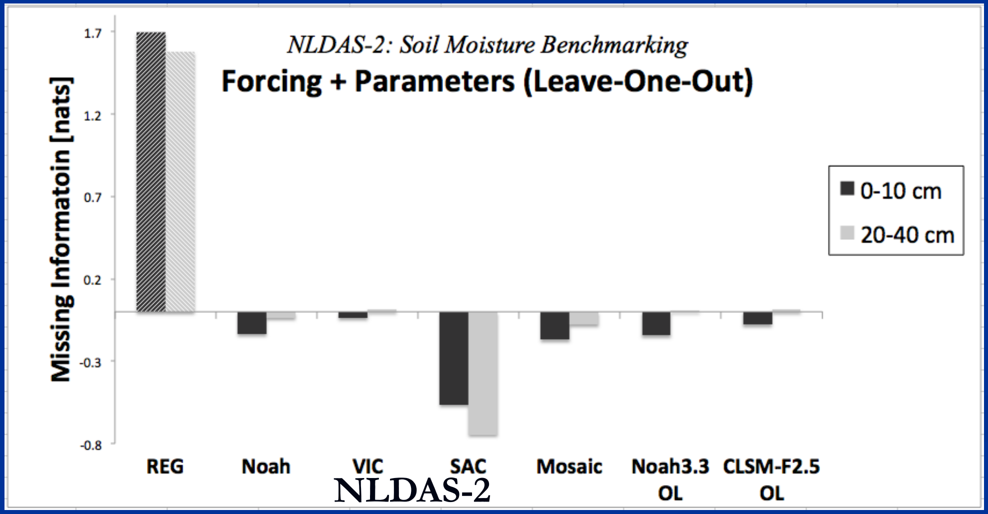


The X-axis – I(X;Y) – is Shannon's mutual information between X and Y; it is equal to the amount of information that Y gives about X and vice versa. Mutual information measures **predictability**. If the Info obtained from the LSM is **Higher** than from the regression model, then the physics from the LSM has **provided additional information**. If the Info obtained from the LSM is **Lower** than from the regression model, then the physics from the LSM has **lost information** through model error.

Regression results

101 SCAN sites with sufficient training data





Standard benchmarking approach (*e.g., Abramowitz, 2005*) using a split data record at each site.

- 15-Day lagged forcing data as inputs
- Single kernel density function trained on ~100 data points at all sites
- Far Left columns measure (Shannon) information missing from Forcings
- Model columns measure <u>additional</u> information lost due to Model + Parameter error

Benchmarking considers info in parameters.

- Single kernel density function trained on data from *external* sites only
- Far Left columns measure (Shannon) information missing from Forcings + Parameters
- Model columns measure info loss (or gain) from model physics.

Discussion

- The previous analyses and description of tools are towards the development of a **systematic approach** to LSM benchmarking.
- LSM parameters should be **distributed and optimized** for the validation datasets that we most trust (and ideally, co-located).
- The benchmarking environment should be automated and consider all aspects of the water and energy balances (a particular LSM may improve its simulation of snow but degrade the fluxes).
- Probabilistic modeling should be done towards reducing the uncertainty from the LSMs with respect to independent validation data, instead of simply showing improvements in error metrics from a single deterministic realization.
- The end goal of this work is to identify and correct model and parameter deficiencies towards improved model fluxes/states.

Future steps for NLDAS

- Evaluate the VIC-4.1.2.l and SAC-HTET-3.5.6/SNOW-17 Open Loop simulations and add/test the effects of DA
- Simultaneous DA of multiple water balance variables (in progress for Noah and CLSM)
- Add effects of irrigation to the NLDAS system (in progress)
- Further evaluations with the regression model
- Drought uncertainty analysis using LIS-OPT/UE subsystem
- Add latest versions of Noah-MP and CLM LSMs into LIS (groundwater, etc.) and run and test over the NLDAS domain
- Transition the latest version of NLDAS using LIS and DA to NOAA/EMC for near real-time operational data production

Summary

- NLDAS is a successful collaboration project that has produced nearly 34 years of hourly 1/8th-degree surface forcing and land-surface model output over CONUS and parts of Canada/Mexico.
- LSMs from NLDAS Phase 2 and from the next phase of NLDAS are being evaluated against observations using LVT, including soil moisture, surface fluxes, snow, and streamflow.
- Generally, the new/upgraded LSMs of Noah-3.3 & CLSM-F2.5 provide better evaluations compared to observations in NLDAS. With DA, these evaluations have been shown to further improve.
- Regression model analysis was performed to show the info lost (or gained) from the model physics and parameters
- LIS-OPT/UE can be used to develop optimized model parameters as well as a measurement of uncertainty of the LSM outputs.

NLDAS & LIS websites

- NLDAS at NASA:

  http://ldas.gsfc.nasa.gov/nldas/
- NLDAS datasets at the NASA GES DISC:

  http://disc.gsfc.nasa.gov/hydrology/
- NLDAS at NOAA/NCEP/EMC:

  http://www.emc.ncep.noaa.gov/mmb/nldas/
- LIS website at NASA:

  http://lis.gsfc.nasa.gov/
- LVT website at NASA:

  http://lis.gsfc.nasa.gov/LVT/