Similarity Assessment of Land Surface Model Outputs in the North American Land Data Assimilation System

Sujay V. Kumar[1], Shugong Wang[1,2],David M. Mocko[1,2], Christa Peters-Lidard[3],

Youlong Xia[4]

[1] *Hydrological Sciences Laboratory, NASA GSFC, Greenbelt, MD*

[2] *SAIC, McLean, VA*

[3] *Earth Science Division, NASA GSFC, Greenbelt, MD*

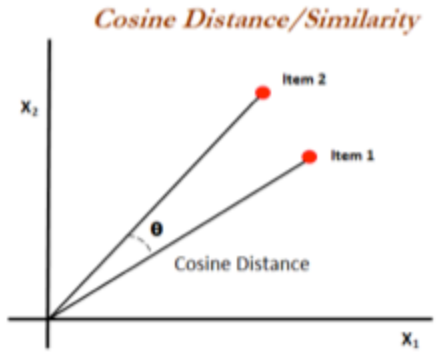
[4] *IMSG, Environmental Modeling Center, NOAA NCEP, College Park, MD*

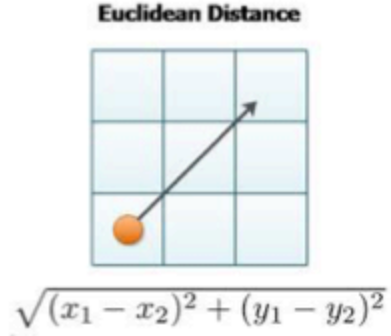
Background

- Multi-model ensembles often used to quantify and understand model uncertainty
- Combining individual model estimates is generally good – leads to increased skill, as individual model errors tend to cancel each other out.
- An evaluation of multimodel ensembles with and without observational constraints is often useful for evaluating the predictability and uncertainty sources of models
- But, observational constraints are not always possible.
- We also assume that the constituent models (and predictions) are independent of each other.

Similarity

- Criterion for assessing things of the same kind
- Information-based
  - Correlation
- Distance-based
  - Mean-squared error
- Considers both commonality and differences





If the constituent models are too similar, they add little additional information to the ensemble.
If the models are too dissimilar, it may be indicative of systematic errors
Assessment of similarity could be a criteria for determining an ensemble.

Methodology

- A latent variable model employing a confirmatory factor analysis (CFA) is used to quantify unmeasured sources of similarity and variability
- A latent variable model is a statistical approach that relates observed variables to "latent"/unobserved variables

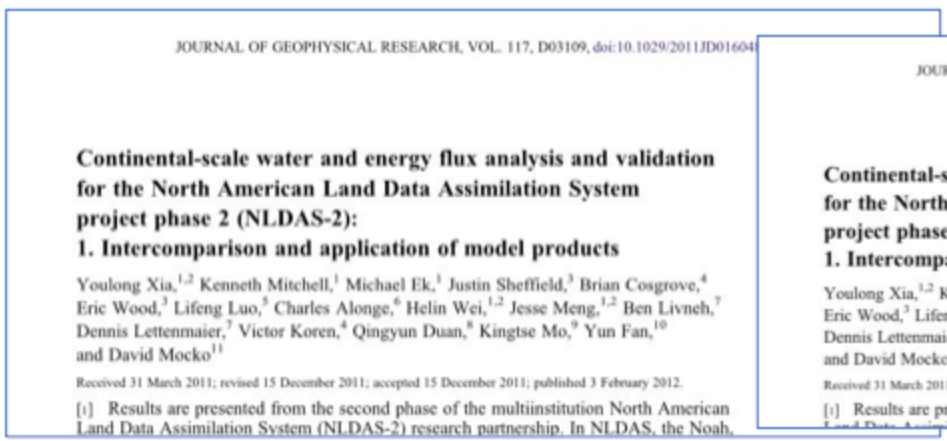
$$x_k = \mu + \lambda_k f + e_k$$

x_k is the modeled estimate for a given variable from model *k;*
μ is constant (to all models);
e_k an independent term related to model *k;*
f is the standardized common factor across the models;
The regression coefficient λ_k is called the **factor loading** for model *k* that describes how strongly the observed variable is associated with the common factor (ranges from -1 to 1)

- Values of λ_k close to -1 or 1 indicate that the common factor strongly affects the variable
  - λ_k close to 1 indicates that the model output is close to the common factor
  - λ_k close to -1 indicates that the model output has a trend opposite to that of the common factor
- λ_k close to zero indicate that the common factor has a weak effect on the ensemble

Approach

- The similarity evaluations are conducted using 10 year (2002-2012) outputs from a suite of LSMs in the NLDAS project.
  - 4 operational LSMs (Noah28, Mosaic, VIC403, SAC)
  - 4 'new' LSMs (Noah36, CLSM, VIC412,Noah-MP)
  - Comparing anomalies of latent, sensible heat fluxes, root zone soil moisture, snow water equivalent and terrestrial water storage.
- NLDAS-1 papers documented deficiencies in model formulations, leading to the NLDAS-2 phase.
- NLDAS-2 evaluations show a greater level of agreement between the constituent models (compared to NLDAS-1).

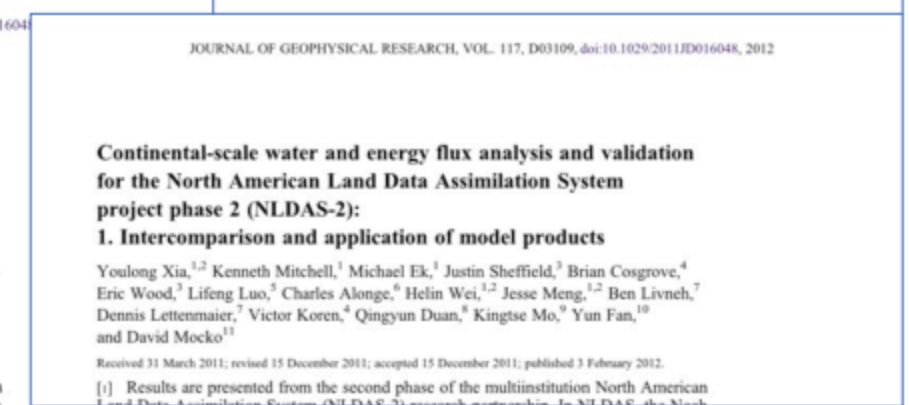

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 117, D03109, doi:10.1029/2011JD016048

Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products

Youlong Xia,[1,2] Kenneth Mitchell,[1] Michael Ek,[1] Justin Sheffield,[3] Brian Cosgrove,[4] Eric Wood,[3] Lifeng Luo,[5] Charles Alonge,[6] Helin Wei,[1,2] Jesse Meng,[1,2] Ben Livneh,[7] Dennis Lettenmaier,[7] Victor Koren,[4] Qingyun Duan,[8] Kingtse Mo,[9] Yun Fan,[10] and David Mocko[11]

Received 31 March 2011; revised 15 December 2011; accepted 15 December 2011; published 3 February 2012.

[1] Results are presented from the second phase of the multiinstitution North American Land Data Assimilation System (NLDAS-2) research partnership. In NLDAS, the Noah,


JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 117, D03109, doi:10.1029/2011JD016048, 2012

Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products

Youlong Xia,[1,2] Kenneth Mitchell,[1] Michael Ek,[1] Justin Sheffield,[3] Brian Cosgrove,[4] Eric Wood,[3] Lifeng Luo,[5] Charles Alonge,[6] Helin Wei,[1,2] Jesse Meng,[1,2] Ben Livneh,[7] Dennis Lettenmaier,[7] Victor Koren,[4] Qingyun Duan,[8] Kingtse Mo,[9] Yun Fan,[10] and David Mocko[11]

Received 31 March 2011; revised 15 December 2011; accepted 15 December 2011; published 3 February 2012.

[1] Results are presented from the second phase of the multiinstitution North American

Comparison and Assessment of Three Advanced Land Surface Models in Simulating Terrestrial Water Storage Components over the United States

YOULONG XIA,[a] DAVID MOCKO,[b] MAOYI HUANG,[c] BAILING LI,[d] MATTHEW RODELL,[b] KENNETH E. MITCHELL,[e] XITIAN CAI,[f] AND MICHAEL B. EK[g]

[a] *I. M. Systems Group at NOAA/NCEP/Environmental Modeling Center, College Park, Maryland*
[b] *Hydrological Science Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland*
[c] *Pacific Northwest National Laboratory, Richland, Washington*
[d] *Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland*
[e] *Prescient Weather Ltd., State College, Pennsylvania*
[f] *Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey*
[g] *NOAA/NCEP/Environmental Modeling Center, College Park, Maryland*

(Manuscript received 13 May 2016, in final form 28 November 2016)

ABSTRACT

To prepare for the next-generation North American Land Data Assimilation System (NLDAS), three advanced land surface models [LSMs; i.e., Community Land Model, version 4.0 (CLM4.0); Noah LSM with

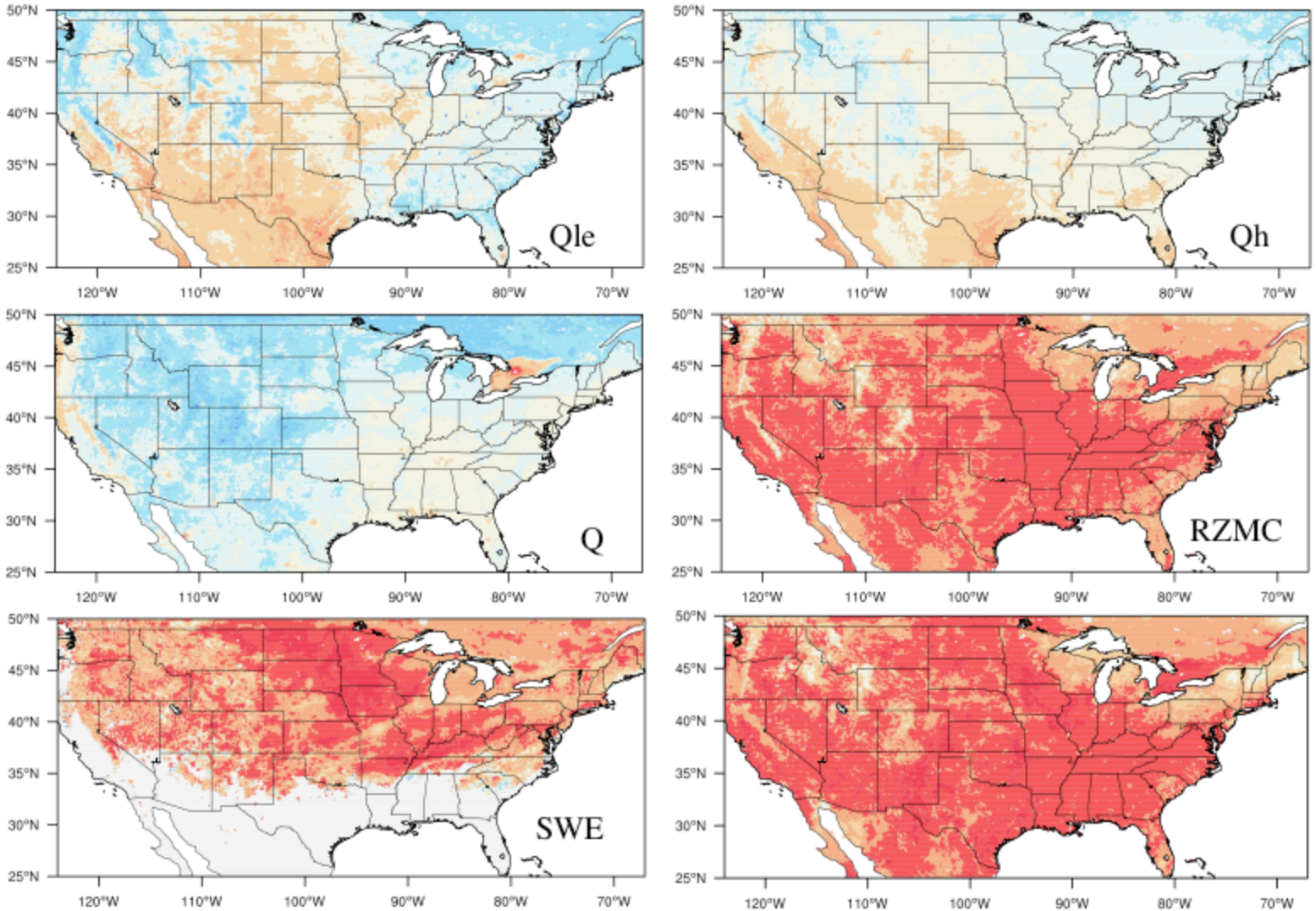
Inter-model similarity assessment

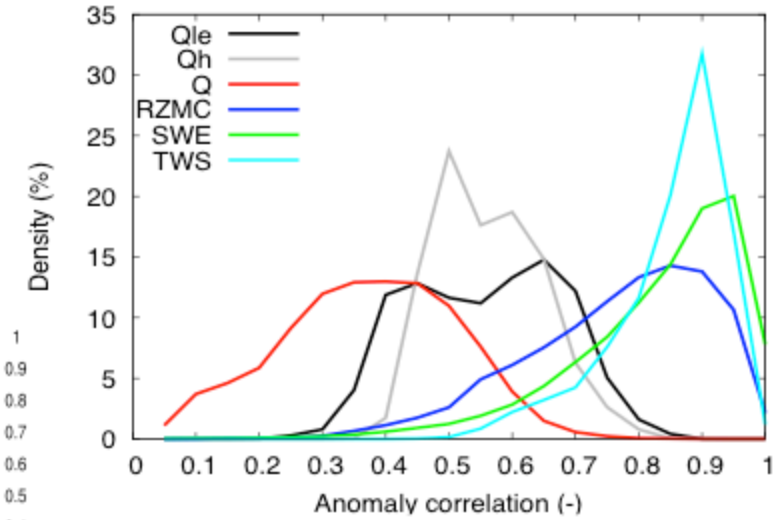
- Average anomaly correlation among the 8 LSMs

Qle	Noah28	Mosaic	VIC403	SAC	Noah36	CLSM	VIC412	NoahMP
Noah28	1.0							
Mosaic		1.0						
VIC403			1.0					
SAC				1.0				
Noah36					1.0			
CLSM						1.0		
VIC412							1.0	
NoahMP								1.0

Average of the off-diagonal elements of the 8x8 matrix, for each variable, representing the inter-model correlations.
A first order estimate of where the models agree with each other and by how much.

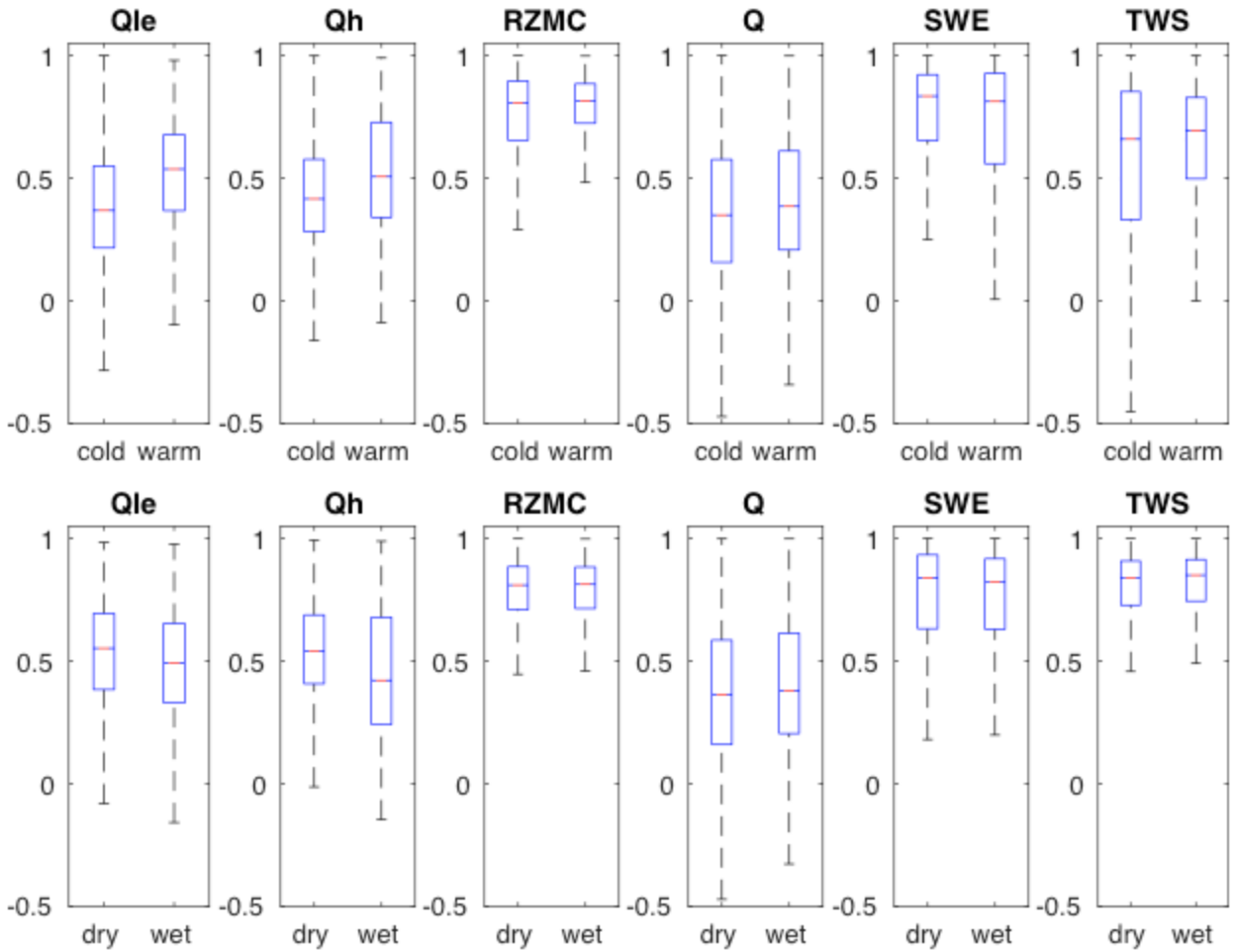
Inter-model similarity (average anomaly R)





- Most dissimilar for Runoff
- Most similar for TWS, RZMC and SWE
- Moderate level of agreement/disagreement for the fluxes

Inter-model similarity (influence of seasonality)



Increased agreement for fluxes, soil moisture, runoff and TWS for warm periods; SWE for cold periods

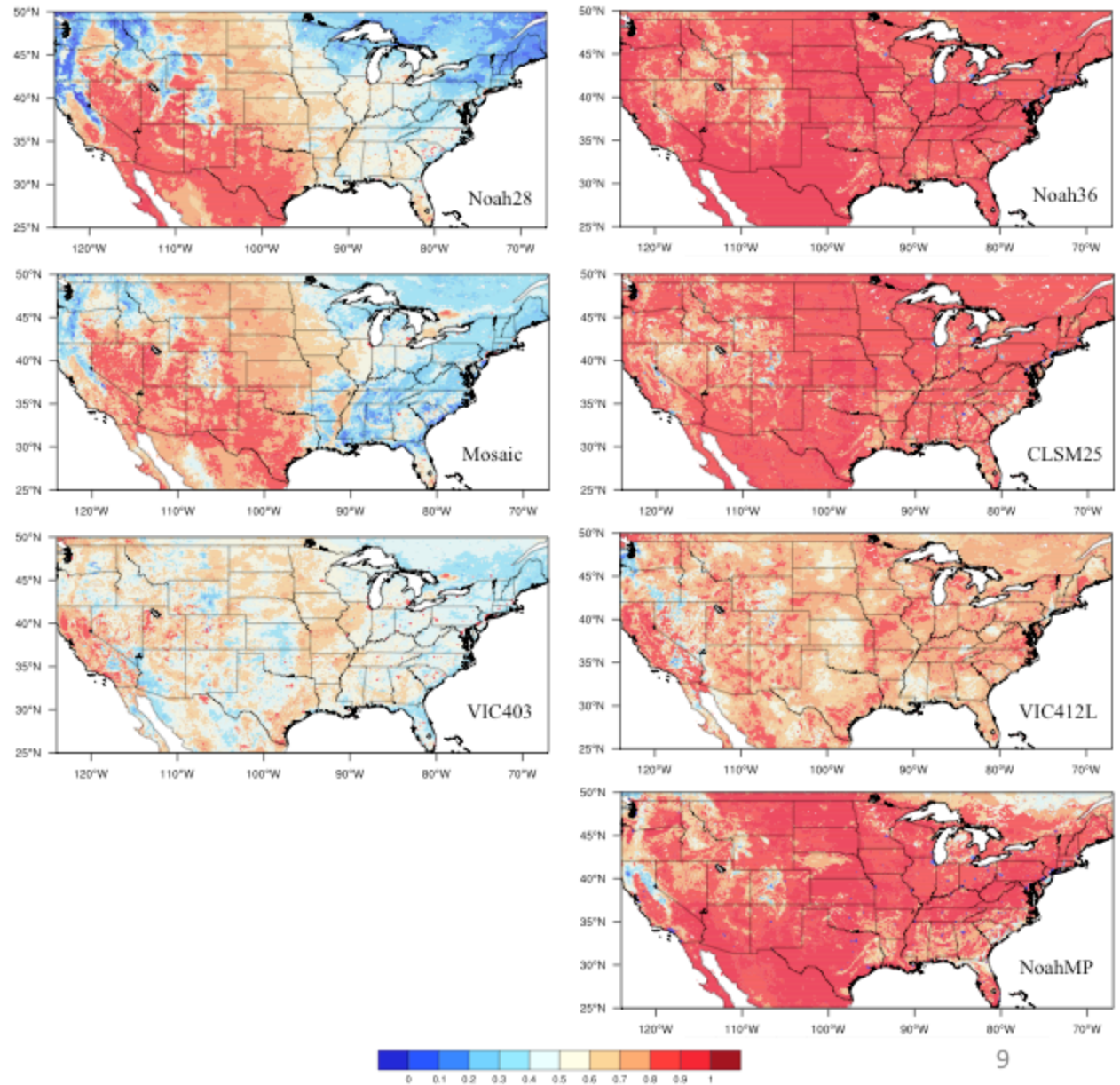
Comparable levels of agreement for both dry and wet days (except for Qh)

Factor loadings (latent heat flux)

Most "new" models have high factor loading values compared to the operational models.

New version of VIC has higher λ_k values, implying that the new version is closer to the common factor

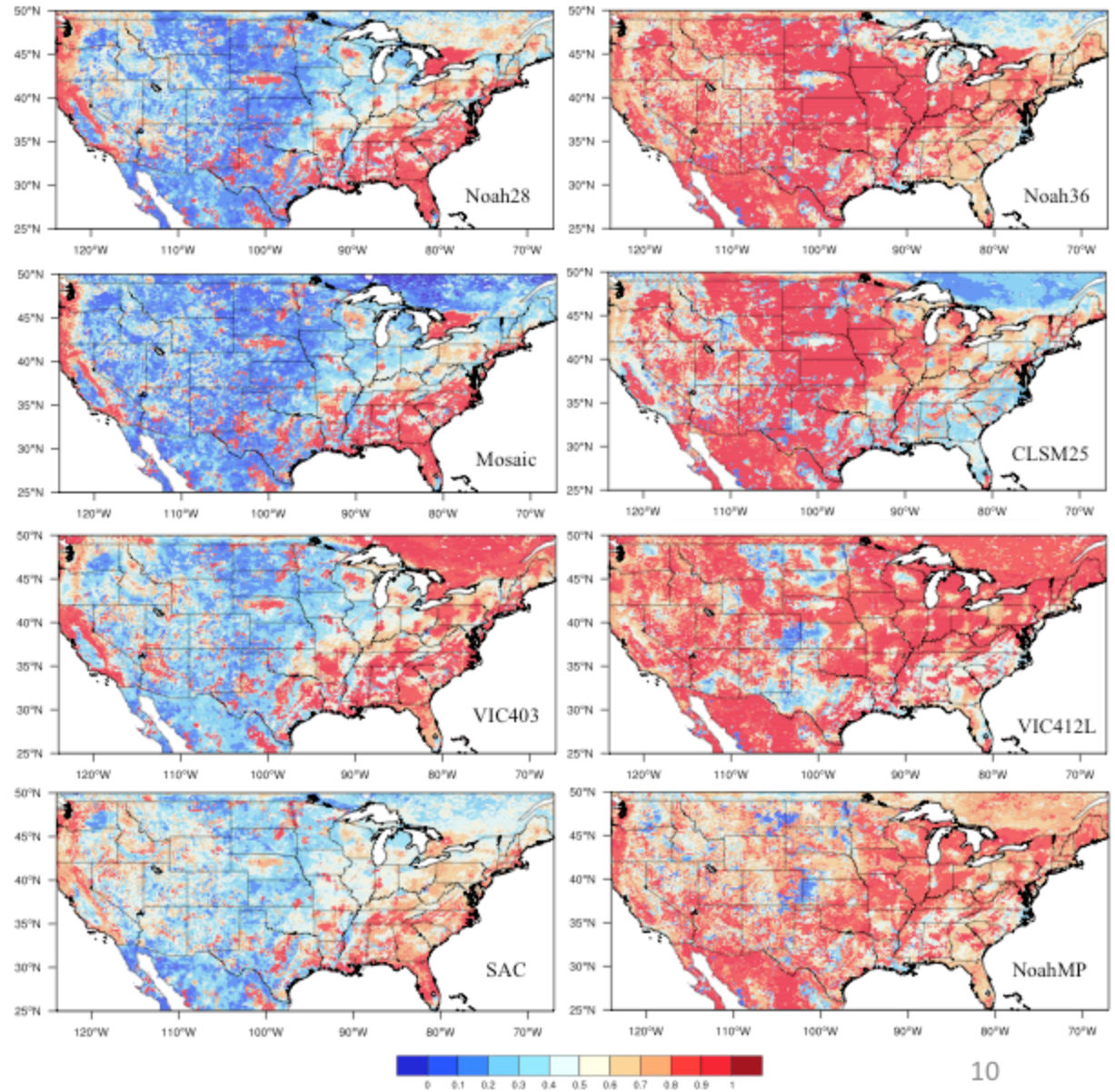
Factor loadings are closer in the water limited domains, differ more in the energy limited domains; more variability over areas with cold season processes.



Factor loadings (runoff)

λ_k values show two distinct groups - new vs old (new models are closer to the common factor)

Older models: Larger λ_k values over the Southeast U.S. and over the West Coast; smaller values over the Central U.S.
New models: Essentially the opposite behavior, though the level of agreement with the common factor is strong

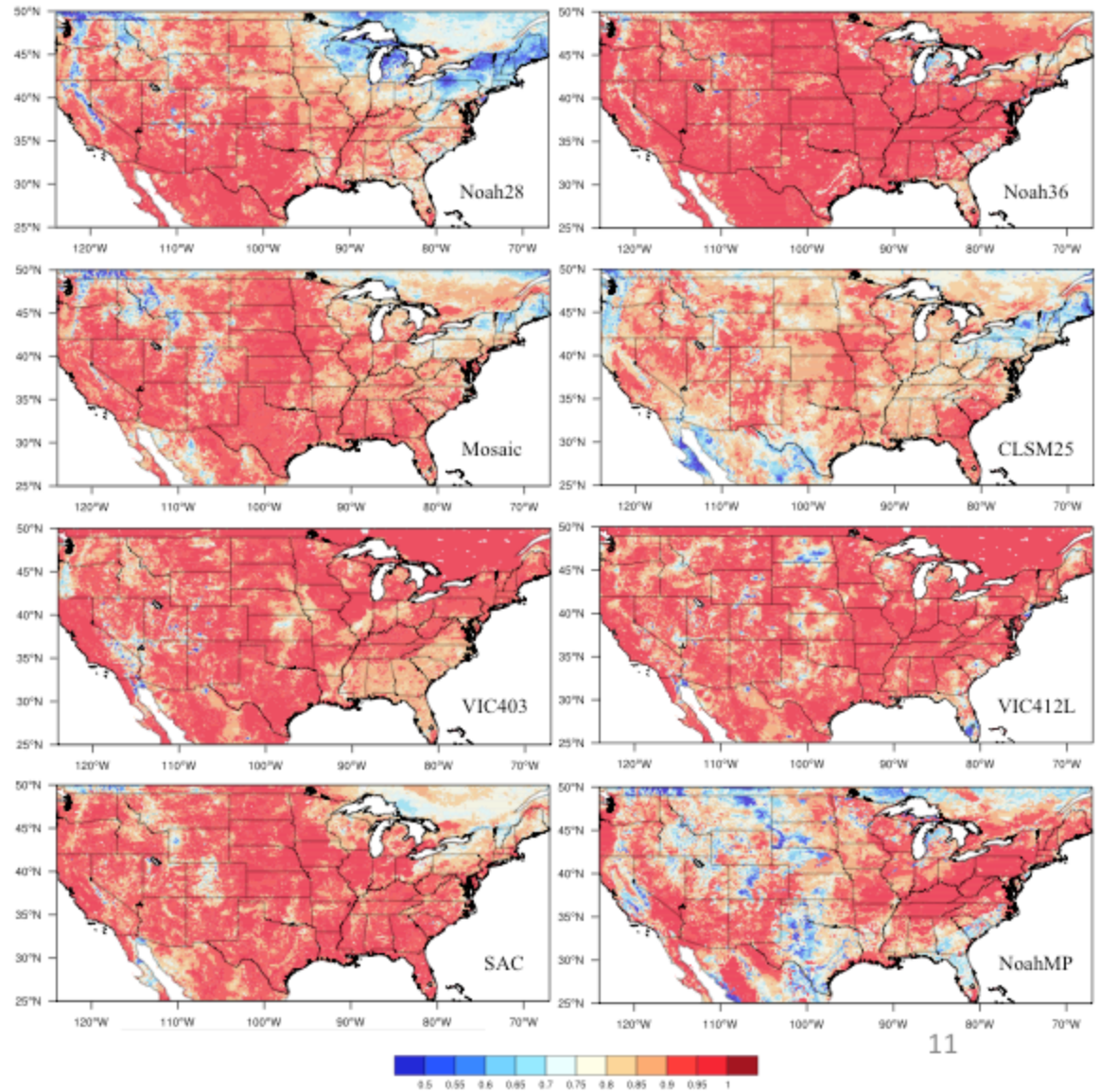


Factor loadings (root zone soil moisture)

All models show strong agreement with the common factor (VIC412 the highest and CLSM the lowest)

The contrast between the new and old models is less.

In the anomaly space, the contribution of an individual model is less (significant for applications such as drought estimation).

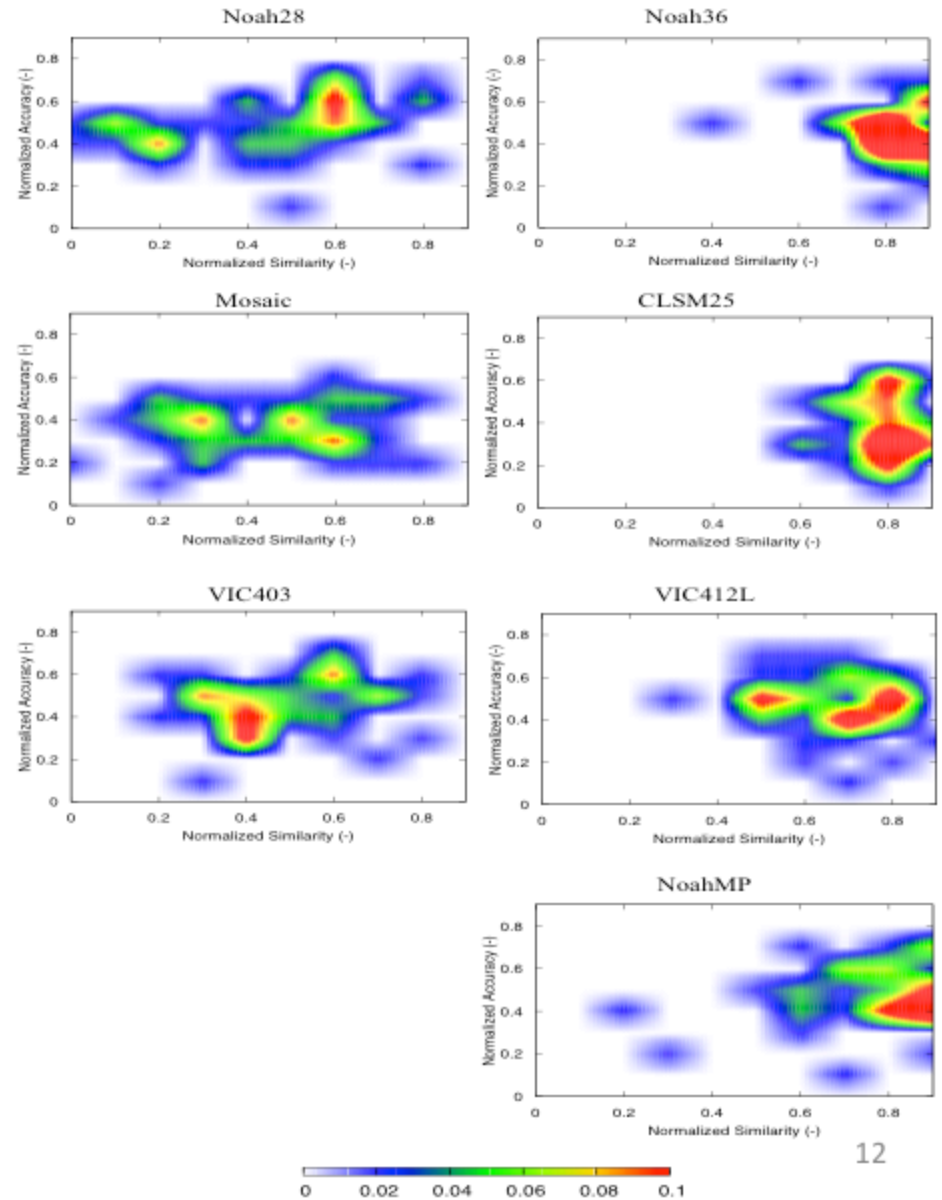


Evaluation of model similarity in relation to accuracy (Qle)

Accuracy is evaluated by comparing the fluxes against Ameriflux in-situ measurements

The older models show considerable spread in the similarity space compared to the newer models.

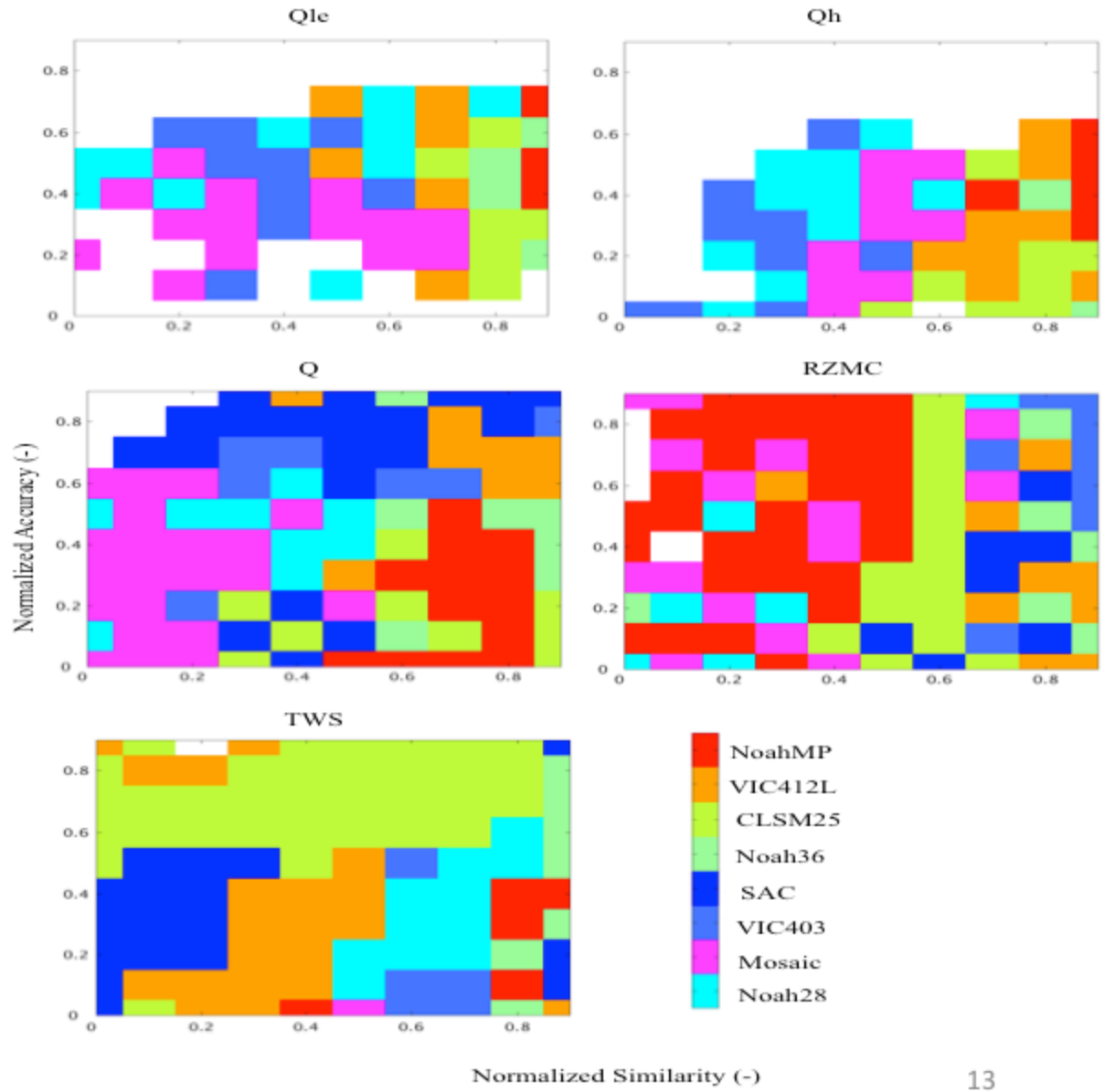
The accuracy is marginally higher in the new models; The new models are more skillful and more similar to each other



Ranking the models in the similarity -accuracy space

Q is evaluated using USGS streamflow data; soil moisture estimates are compared against SCAN in-situ data; TWS compared against GRACE.

Generally, the newer models (warm colors) span the high similarity-high accuracy range.



Summary

- Similarity is a useful criteria to evaluate the models toward their utility to an ensemble, without requiring observational constraints.
- If the models are too similar, their utility to the overall ensemble is low. Conversely, if the constituent models are deficient in their formulations, the model estimates would be very different from each other.
- Similarity analysis applied to the NLDAS models indicate that runoff estimates are most dissimilar; soil moisture, SWE and TWS estimates are very similar.
- The NLDAS operational models showed weaker association with the common factor of the ensemble and the new models showed stronger association with the latent factor.
- The results suggest that model development efforts have pushed the models to be more skillful, but also more similar to each other.

Kumar, S.V., S. Wang, D. Mocko, C.D. Peters-Lidard, Y. Xia (2017): Similarity assessment of land surface model outputs in the North American Land Data Assimilation System (NLDAS), *Water Resources Research*, 53,8941-8965.