Similarity Assessment of NLDAS Multi-Model Ensemble Outputs

Shugong Wang[1,2], Sujay V. Kumar[1], David M. Mocko[1,2], Christa D. Peters-Lidard[1], Youlong Xia[3,4], Michael B. Ek[3]

1 - NASA Goddard Space Flight Center (GSFC), Greenbelt, MD; 2 - Science Applications International Corporation (SAIC) at Greenbelt, MD
3 - Environmental Modeling Center, NOAA/NCEP, College Park, MD; 4 - I.M. Systems Group (IMSG) at College Park, MD

1. Introduction

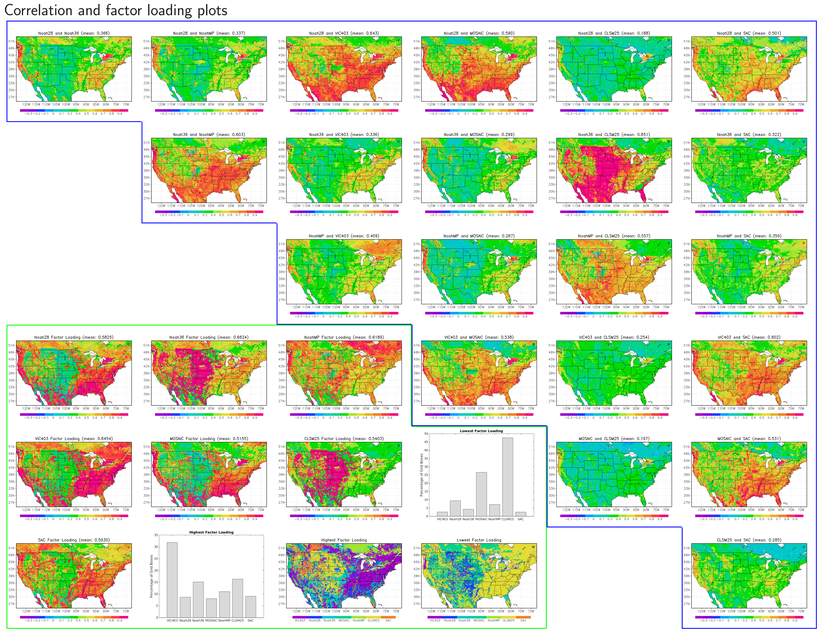
- Multi-model ensembles have been used in weather, climate, and hydrologic projects to produce ensemble mean estimates.
- The North American Land Data Assimilation System, NLDAS, uses a multi-model ensemble to produce land surface states, energy, and water fluxes.
- The current NLDAS models yield different simulations because of different development philosophies, formulations, and parameterizations applied.
- The ensemble members have different levels of similarity and dissimilarity depending on factors such as climatic regimes, seasonality, model parameters, topography, and geography.
  - If multi-model outputs are very similar, there might be little additional information to the multi-model ensemble.
  - If one model is very dissimilar, it may indicate errors in the simulation, leading to increase in the mean bias and the variance of the ensemble.
- A similarity assessment of NLDAS multi-model ensemble outputs is necessary for assessing the fidelity and usefulness of the ensemble.
  - The assessment is not against observations but the member importance to ensemble.

2. Assessment Design

- Seven models are included.
  - The NLDAS-2 models: VIC-4.0.3, Noah-2.8, MOSAIC, and SAC.
  - Possible next phrase NLDAS models: Noah-3.6, Noah-MP, Catchment model (CLSM-F2.5).
    - Noah-MP (version 3.6) is configured with dynamic vegetation.
- Six output variables are assessed: Runoff (Qs+Qsb), SWE, root one soil moisture (SMC100) content, latent heat flux (Qle), sensible heat flux (Qh), and ground heat flux (Qg).
  - SAC is excluded from the analysis of Qh, Qle, and Qg.
- Eleven years of model outputs are used (2002-2012, daily average for SWE, SMC100, Qle, Qh, and Qg, daily sum of total runoff).
- Two similarity/dissimilarity metrics
  - Correlation coefficient - correlation of output time series among the models
  - Factor loading - factor analysis with single common factor $\mathbf{x} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f} + \mathbf{e}$.
    - Loadings can range from -1 to 1.
    - Loadings close to -1 or 1 indicate that the factor strongly affects the variable.
    - Loadings close to zero indicate that the factor has a weak effect on the variable.
    - MATLAB `factoran` is used to compute the maximum likelihood estimate of the factor loadings.
    - Reference: Christensen & Sain, 2012
- NLDAS domain: CONUS + Canada (up to 53° North) + Mexico (down to 25° North)
  - The CONUS area has better parameters and forcing data.

3.1 Results of Total Runoff (Runoff = Qs + Qsb)

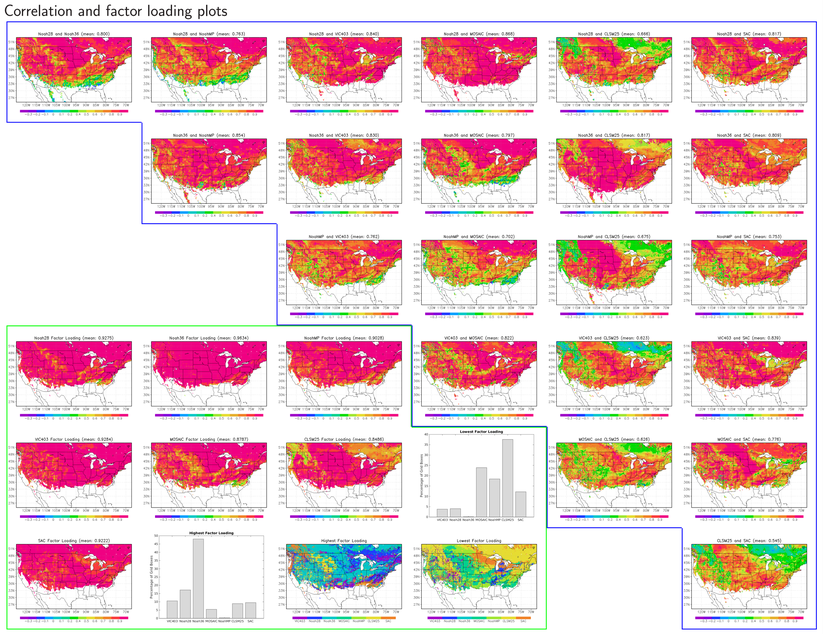
- Correlation and factor loading plots



- Total runoff has the lowest similarities compared with other variables. The NLDAS-2 models are more similar to each other than the rest of combinations, except Noah-36 and CLSM-F2.5. Over 30% of grid boxes, VIC-4.0.3 is most similar to other models. In nearly 50% grid boxes (in the eastern part of the NLDAS domain), CLSM-F25 is most dissimilar to other models. The similarities among most of models are distinctively different along the US-Canada boundary because of the quality of forcing data and parameters. In general, the Rocky Mountains area has lower similarities compared with the rest areas.

3.2 Results of Snow Water Equivalent (SWE)

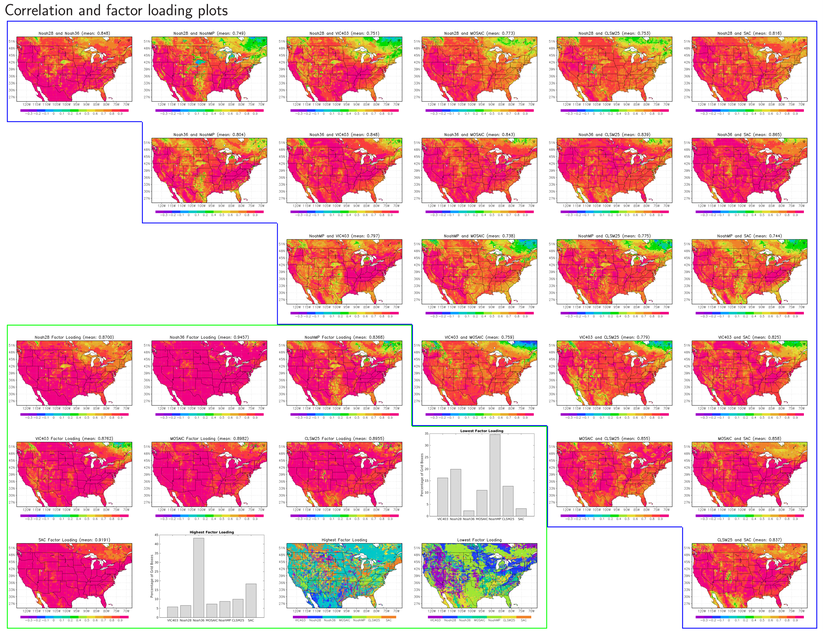
- Correlation and factor loading plots



- The SWE analysis only includes data in snow seasons. The similarities are usually higher in the Great Plains than the rest areas and lower in the Rocky Mountains area, the eastern Canada and the southern US. In about 50% snow-covered grid boxes, Noah-36 is most similar to other models. In nearly 40% grid boxes (mostly in Canada), CLSM-F25 is most dissimilar to others.

3.3 Results of Root Zone Soil Moisture (0-100cm)

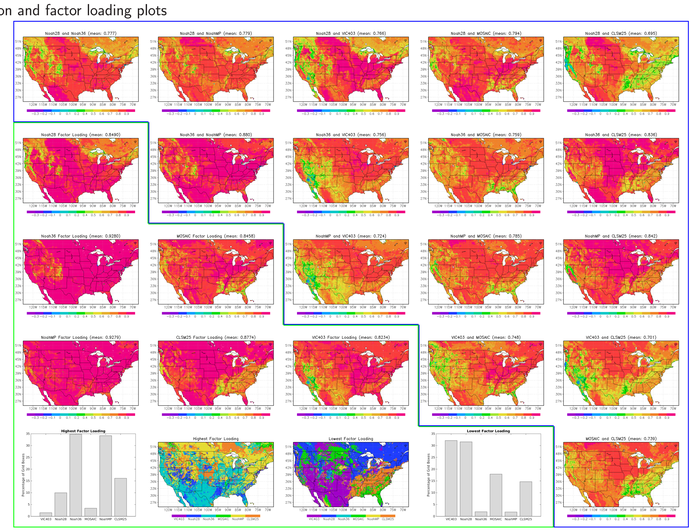
- Correlation and factor loading plots



- Overall, the similarities of root zone soil moisture do not show a strong dependence on location within the CONUS area. Individually, model outputs may have higher similarity in different areas. But the similarities among all models are significantly lower in the Canada part of the NLDAS domain, especially the northeastern corner. In about 45% grid boxes, Noah-36 is most similar to other models. On the other hand, Noah-MP is most dissimilar to other models for nearly 35% grid boxes.

3.4 Results of Latent Heat Flux (Qle)

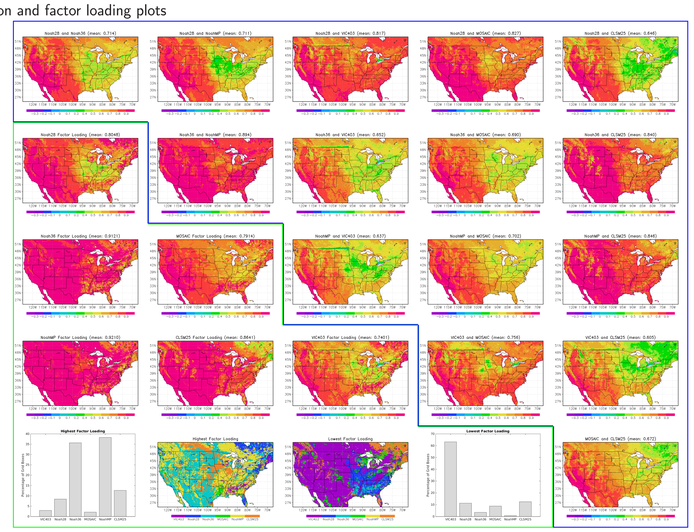
- Correlation and factor loading plots



- All models have higher similarities over flat areas (e.g. the Great Plains) than mountain areas (e.g. the Rocky Mountains and the Appalachian Mountains). In nearly 35% of grid boxes, Noah-3.6 and Noah-MP are most similar to the rest of models. In over 30% grid boxes, VIC-4.0.3 and Noah-2.8 are most dissimilar to other models.

3.5 Sensible Heat Flux (Qh)

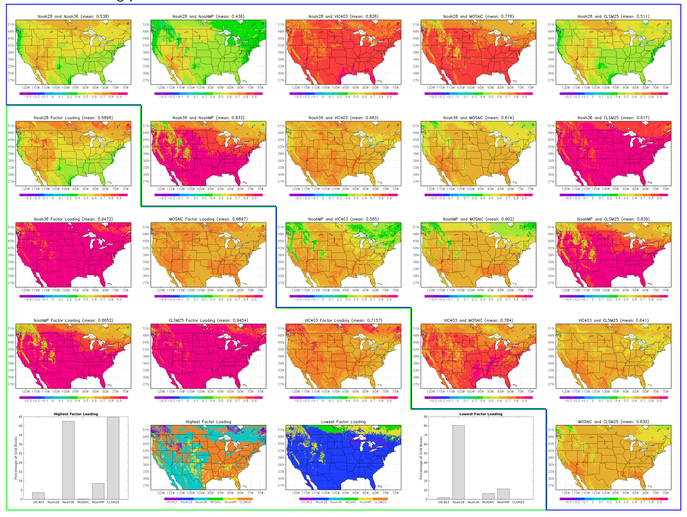
- Correlation and factor loading plots



- The similarities of sensible heat fluxes have a strong location dependence for all models. There are higher similarities in the western states of US and lower similarities within the Mississippi River Basin and northeastern corner of the NLDAS domain. This may relate to dry or wet conditions of land surface. Respectively, Noah-MP and Noah-36 are most similar to other models in over 35% grid boxes. VIC-4.0.3 is most dissimilar to other models in over 60% grid boxes.

3.6 Ground Heat Flux (Qg)

- Correlation and factor loading plots



- The similarities of ground heat flux are different significantly among models. Noah-3.6, Noah-MP are CLSM-F25 very similar within the CONUS area and less similar in Canada and the Rock Mountains area. The NLDAS-2 models are relatively similar to each other. Noah-28 is much less similar to Noah-36, Noah-MP and CLSM25. Respectively, CLSM-F25 and Noah-MP are most similar to other models in over 40% grid boxes. In about 80% grid boxes, Noah-28 is most dissimilar to other models.

4. Conclusions

- The model output similarities depends on many factors, including but not limited to model physics, topography, and qualities of forcing data and parameters.
  - Different model physics and parameterization yield different simulation skills, which causes outputs of one model to have different similarities to those of other models. For example, VIC 4.0.3 is most similar to other models in terms of total runoff but it is most dissimilar to other models in terms of sensible heat flux (Qh).
  - Most model outputs have higher similarities within the CONUS area than the Canadian and the Mexican parts of the NLDAS domain because there are better forcing data and parameters in CONUS.
  - In general, model outputs have lower similarities over mountain areas with steep topography, such as the Rocky Mountains.
  - The value of adding a model to an ensemble depends on variables and locations. For the same model, variables usually contribute to the ensemble differently.
- Correlation coefficient and factor loading are consistent in evaluating the similarity and dissimilarity of model outputs.
  - A higher factor loading indicates that the output of a model has higher correlations with the outputs of other models.
  - Factor loading is effective to assess the model contribution to ensemble. If an ensemble mean is a weighted average, the model output with higher factor loading contributes to it more than those with lower factor loadings.
  - It is essential to establish correlation and factor loading thresholds for deciding the significance of adding a model to an ensemble. Only models which are not very similar or dissimilar to others may be added to the ensemble.
- The assessment study reveals that the new models have a bigger contribution to the ensemble than the NLDAS-2 models with the exception of total runoff. For the largest percentage of grid boxes,
  - Noah-3.6 is most similar to the rest of models in SWE and root zone soil moisture;
  - Noah-MP and Noah-3.6 are most similar to the rest of models in sensible heat fluxes and latent heat flux;
  - CLSM-F2.5 and Noah-3.6 are most similar to the rest of models in ground heat flux;
  - VIC-4.0.3 and Noah-2.8 are most dissimilar to the rest of models in latent heat flux;
  - VIC-4.0.3 is most dissimilar to the rest of models in sensible heat flux;
  - Noah-2.8 is most dissimilar to the rest of models in ground heat flux.