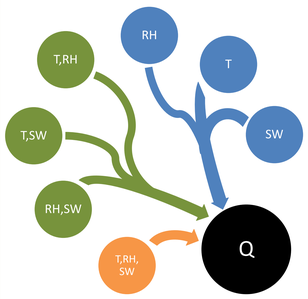
A process network based approach to model intercomparison using SUMMA ensembles

Andrew Bennett[1], Bart Nijssen[1], Martyn Clark[2], and Grey Nearing[3]
[1]University of Washington, Seattle, Washington; [2]University of Saskatchewan; [3]University of Alabama, Tuscaloosa, Alabama

EGU2019-12070

Background

Process networks & model benchmarking



Compute process networks for all combinations of effect of temperature (**T**), relative humidity (**RH**), and shortwave (**SW**) to turbulent heat fluxes (**Q**) using conditional mutual information (**I(X;Y|Z)**) to understand process interconnections

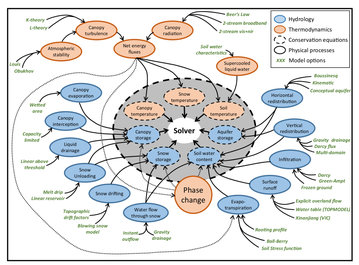
We compare an ensemble of simulations against observations and a statistical model benchmark for both performance and process representation

SUMMA

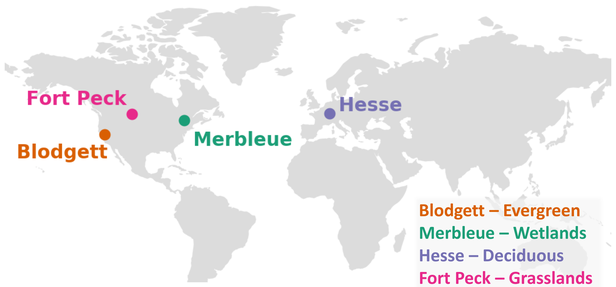
Framework for implementing hydrologic models

User can choose spatial discretization and flux parameterizations

Ensembles can be built in a controlled fashion



Simulations

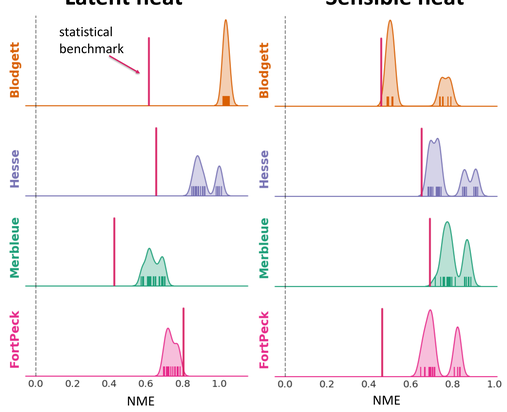


288 Model runs per site:

- 2 parameterizations of vegetation dynamics
- 2 parameterizations of canopy emissivity
- 12 parameterizations of stomatal resistance
- 2 parameterizations of canopy interception
- 3 parameterizations of canopy shortwave

Results

How does our ensemble perform compared to the benchmark?

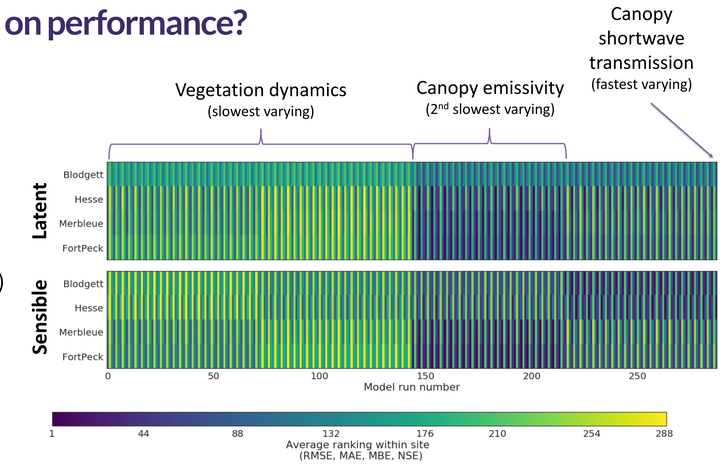


Show distributions of normalized mean error (**NME**) for both heat fluxes

Generally perform better for sensible heat than latent heat across sites

Only outperform the benchmark in latent heat for Fort Peck

What is the effect of each model decision on performance?



Compute multiple error metrics on each ensemble member and rank them (lower rank=better simulation)

Blodgett/Merbleue and Hesse/Fort Peck shared best ranked configurations

All best model configurations used same vegetation dynamics parameterization

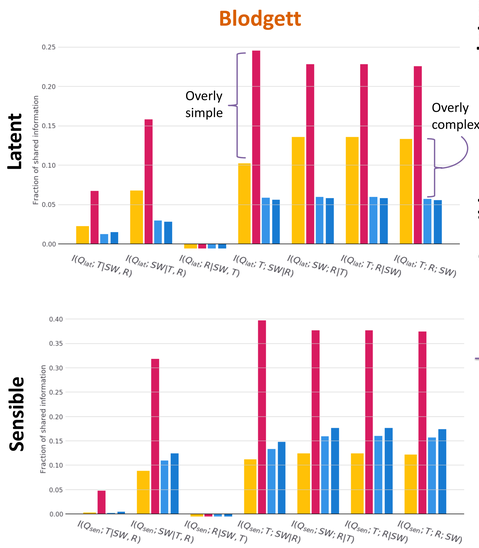
How much information from forcing variables is shared by turbulent heat fluxes?

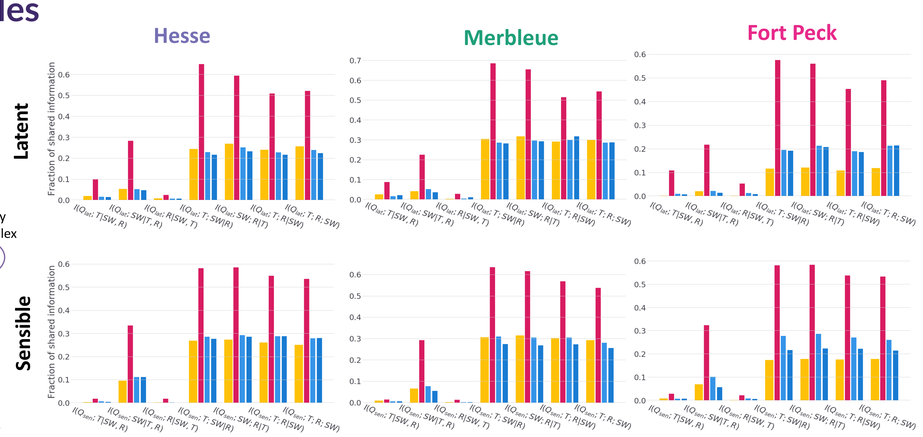
Statistical model oversimplifies connection between forcings and turbulent heat fluxes across all sites

Best performing runs tend to more closely match observed mutual information than worst performing runs

Shortwave is the best single predictor for both latent and sensible heat

Using multiple variables increases mutual information





Blodgett shows overly complex relations for latent heat, also worst performance relative to benchmark

Merbleue and Hesse show close resemblance in marginal structures for both best and worst performing simulations

Cannot distinguish between clusters on performance distributions

Fort Peck shows too simplistic marginal structure across all models for triadic relations

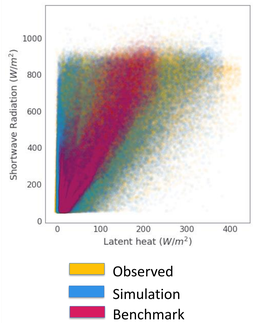
Worse performance has closer representation of marginal structure in Fort Peck

Observed | Benchmark | "Best" simulation | "Worst" simulation

Discussion

Statistical model relies too heavily on shortwave radiation

This oversimplification is seen in scatter plot of data from all sites (benchmark doesn't cover the space well)



Physically based models performance is highly variable across parameterization and site

Future work: How can we maximize both performance and process representation?

Conclusions

Statistical model performs best on error metrics, but are over-constrained by shortwave radiation

Physically based models tend to better match observed shared information between variables but have wide range of performance

This work was supported in part by grants NNX15AI67G and 80NSSC17K0541 from the NASA AIST program.