Estimation of Structural Error in the Community Land Model Using Latent Heat Observations

J. Ray, M. Huang, Z. Hou and L. Swiler

Contact: *jairay@sandia.gov*

Sandia National Laboratories, CA & NM
and
Pacific Northwest National Laboratory.

SAND2014-2431C

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Introduction (1/2)

- Aim: Calibrate the Community Land Model (CLM) using measurements of Latent Heat (LH) at 2 sites
  - Estimate 3 hydrological parameters to which LH is sensitive, along with uncertainty
  - Estimate structural error (model-form error) for LH using 2 different models
- Challenges
  - CLM is expensive – 45 minutes/invocation per site
  - No previous work on optimal site-specific parameters or relative importance of structural versus parametric uncertainty
  - No idea of the ‘shape’ of the structural error
  - No previous work on the Bayesian calibration of CLM parameters (except 1 paper)

Introduction (2/2)

- Our approach to Bayesian calibration of CLM
  - Create surrogates of CLM
  - MCMC calibration of the parameters (3 CLM parameters + those of structural error model)
  - Shortlist physical processes that may be causing the structural error
- So, what is CLM?
  - A model for biogeochemical processes on land
    - Used in conjunction with an atmosphere and ocean model in earth system / climate change simulations
  - Models heat and mass transfer on and under the ground
  - A collection of 1D PDEs (depth), coupled by algebraic equations horizontally
  - Can be forced with measured meteorology if no atmosphere model
  - Soil, vegetation etc. parameters are inputs
  - Can be run in site-specific or global mode

Sandia National Laboratories

Our approach (1/2)

- Pick 2 sites: US-MOz and US-ARM
  - Monthly averaged LH measurements & meteorology available for 48 months
  - 3 most important hydrological parameters are in literature
- Construct surrogate models $\log(LH) = G(p_1, p_2, p_3)$
- Calibrate (estimate parameters) using least-square minimization (“optimal parameters”, $\mathbf{p}_{opt}$)
- Compute data – model mismatch; propose structural error model
- Calibrate CLM parameters and structural error parameters using MCMC (i.e., parameter estimates and uncertainty bounds)

Sandia National Laboratories

Our approach (2/2)

- To investigate
  - Does posterior PDF contain $\mathbf{p}_{opt}$?
  - Does the choice of structural error model affect parameters' PDFs?
  - Does climatological averaging affect parameters' PDFs?
- Constructing surrogate model
  - Model - a 5th order polynomial

$$\log(LH) = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=0}^{5} \sum_{l=0}^{5} a_{ij} p_i^k p_j^l, \quad (k+l) \leq 5$$

  - $\{p_1, p_2, p_3\} = \{F_{drai}, \log(Q_{dm}), b\}$ for US-ARM; $\{S_y\}$ for US-MOz
  - Uniform priors for all parameters
  - Generate a training set of data to capture the parametric dependence of log(LH) - 282 ***p*** samples via QMC sampling
    - Each ***p*** sample -> CLM -> 48-month time-series of log(LH) predictions
  - Construct a separate surrogate model for each month

Sandia National Laboratories

Preventing overfitting of surrogate models

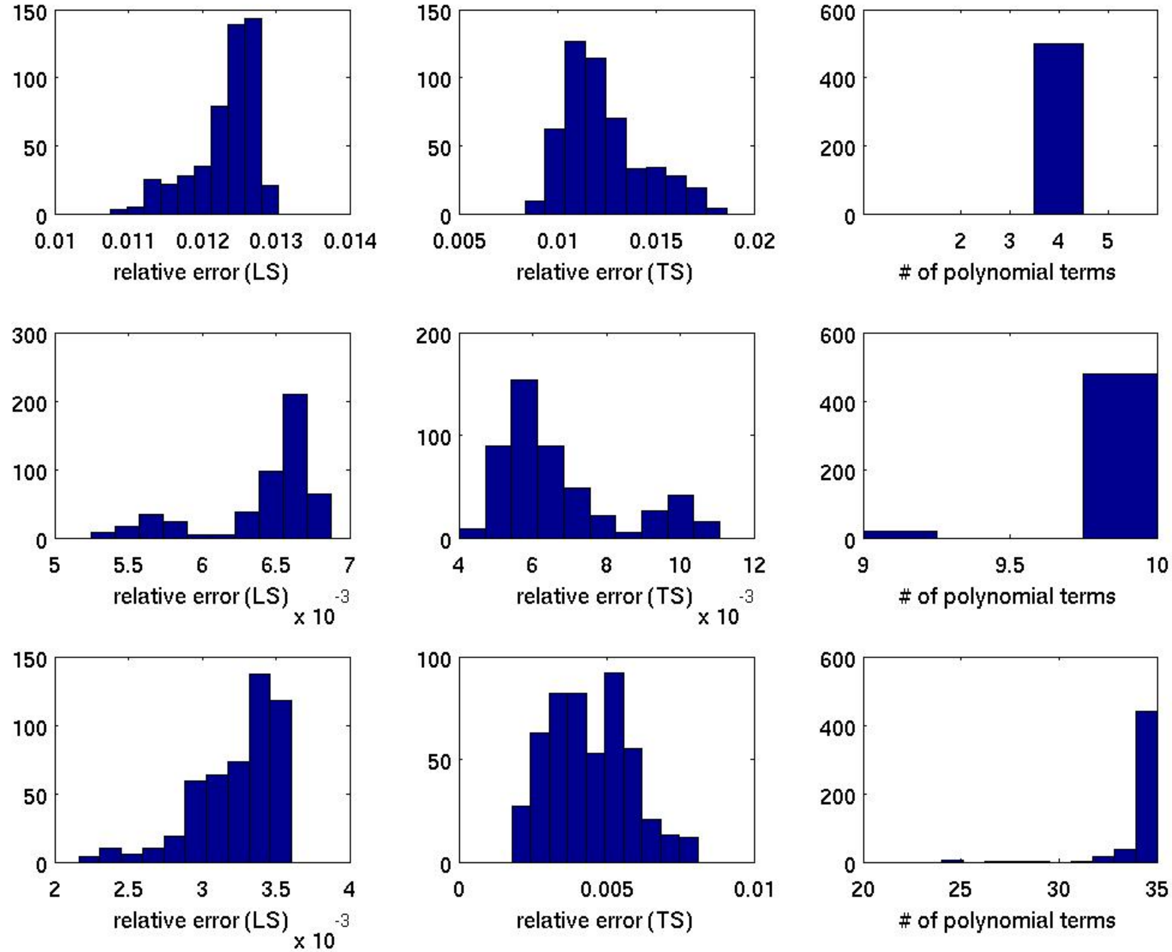
- We need the sparsest polynomial model i.e., set as many a_{ij} to zero as possible
  - Done via AIC and 500-fold cross-validation
- Cross-validation
  - The training set is split into a Learning Set (LS; 85% runs) and a Testing Set (TS)
  - Polynomial model fitted to LS runs and simplified using AIC
    - most cubic and higher-order terms drop off
  - Fitted model is used to predict TS log(LH) values; relative errors are computed for both LS and TS
  - We repeat 500 times using different LS/TS partitions
  - We want models to be equally prediction for LS and TS
    - If overfitting, LS errors < TS errors
  - We also want models to have rel. errors < 10%

Surrogate modeling and calibration

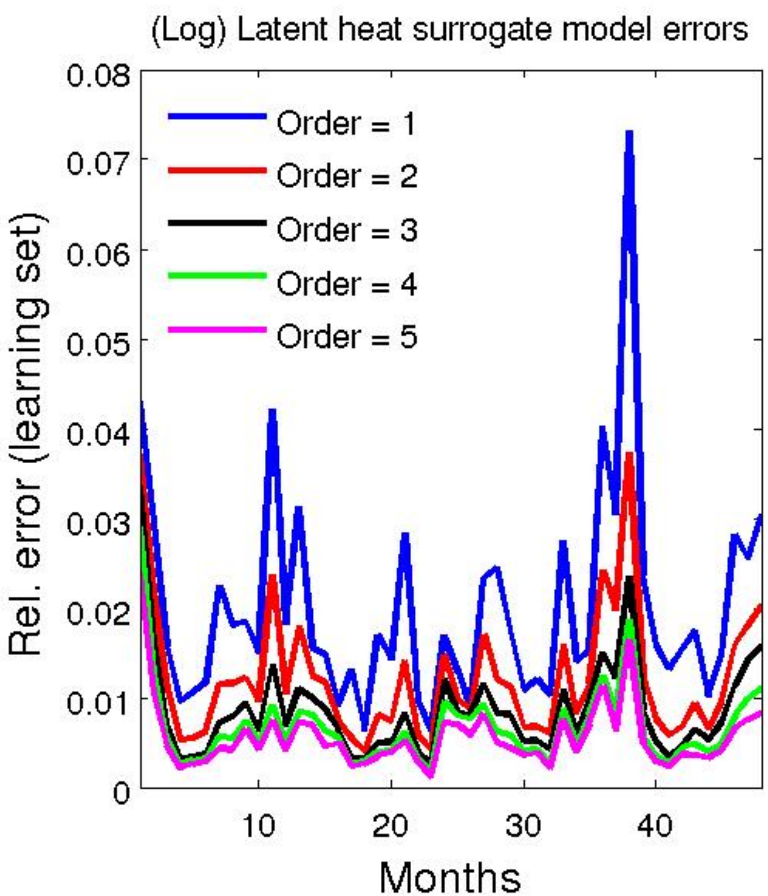
US-ARM

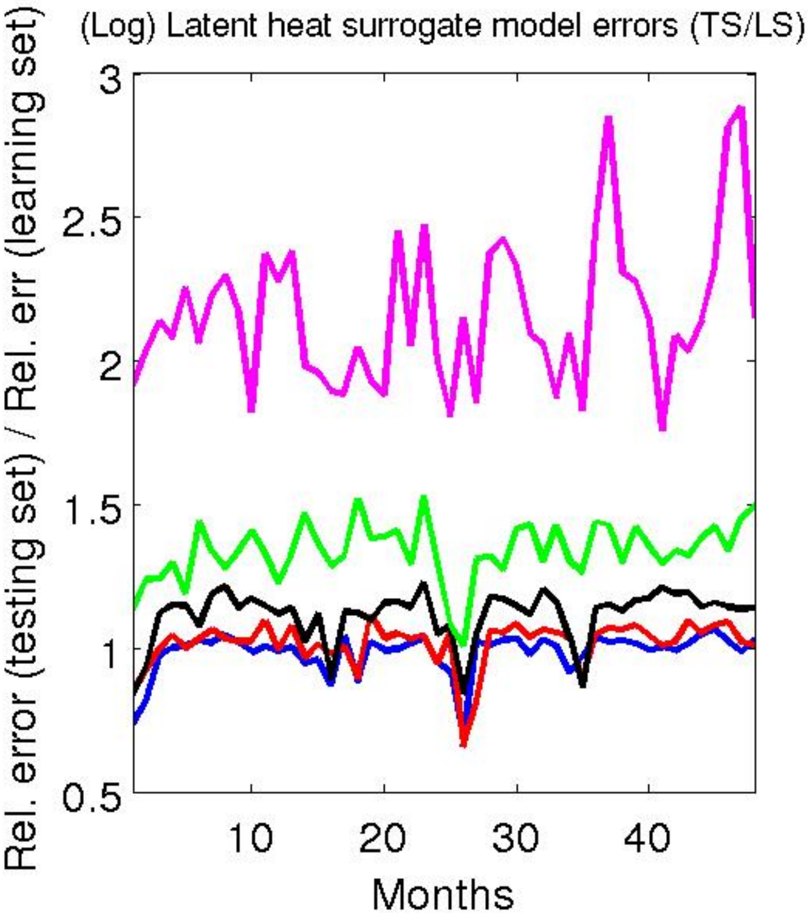
Surrogate model performance – April 2003

- Quartic model suffers from overfitting
- Even the # of terms in polynomial does not remain constant



Surrogate model performance – 48 months

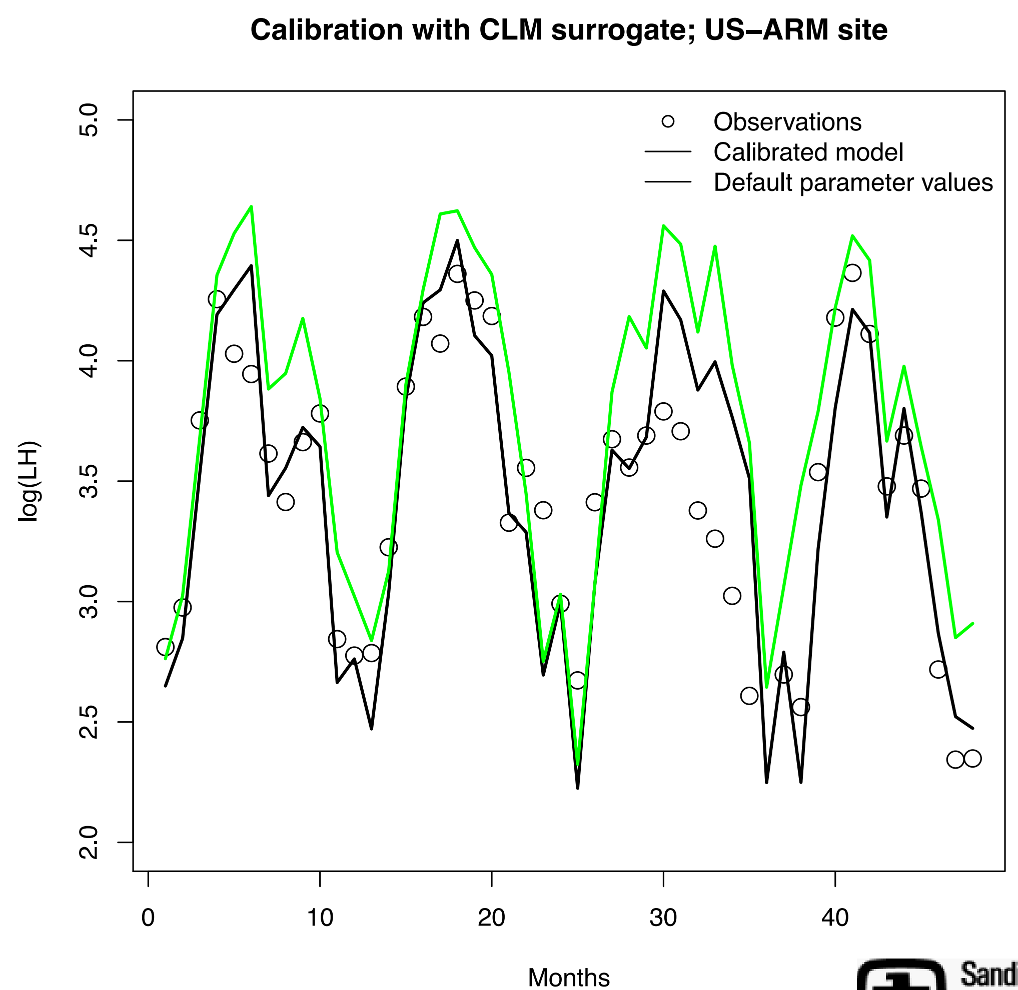




- We will stay with quadratic models

Deterministic calibration

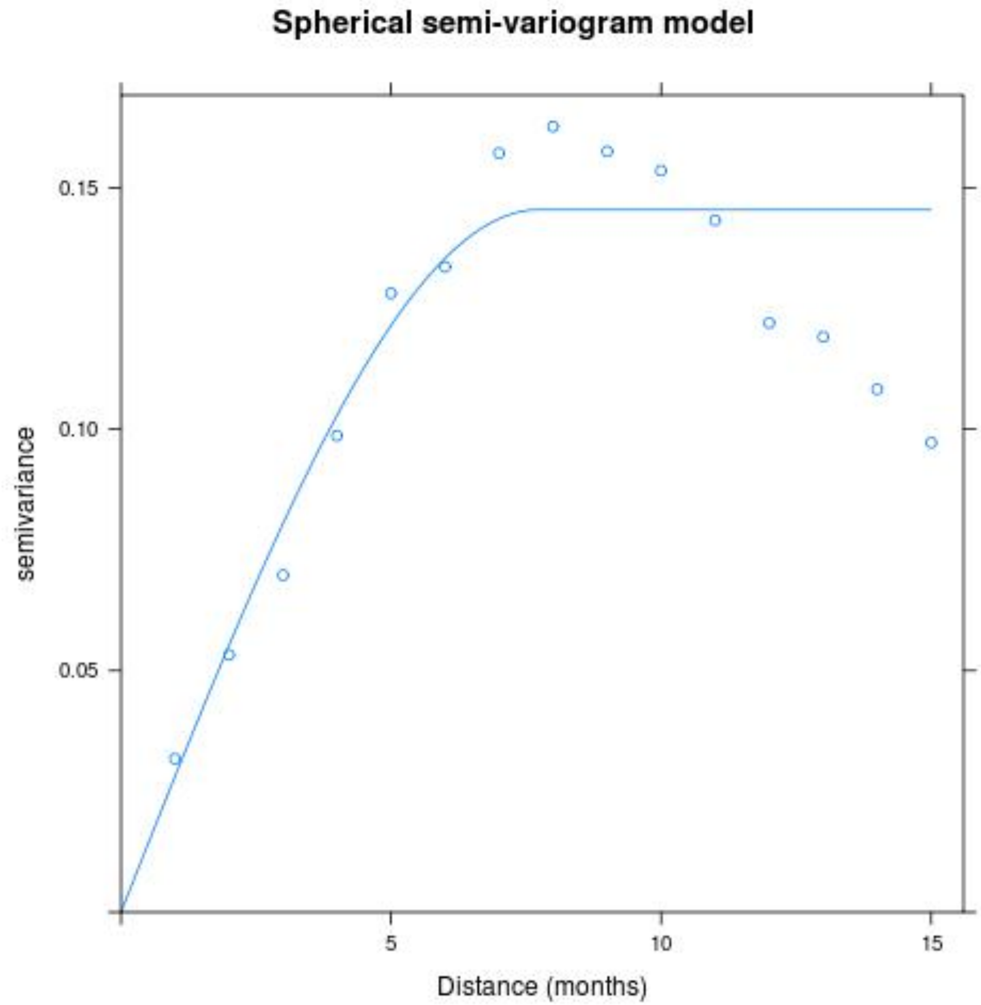
- Least squares fit using L-BFGS-B
  - Trouble; multiple minima
- $\mathbf{p}_{opt}$ = {F_{drai}, log(Q_{dm}), b} = {0.97, log(10^{-2}), 0.1}
- Improvement in predictive skill
- Suggests 2 structural error models
  - uncorrelated errors i.e., $\varepsilon \sim N(0, \sigma^2)$
  - Temporally correlated errors, i.e. $\varepsilon \sim N(0, \Gamma)$



Sandia National Laboratories

Structural error model

- Uncorrelated errors
  - i.i.d. Gaussians
  - Prior model for σ^2 is an inverse Gamma
- Correlated error model
  - Modeled as a stationary 1D multiGaussian field
  - Spherical semi-variogram for covariance
    - Sill (σ^2) and range (τ) are to be estimated from data
    - Exponential priors for sill and range
    - Prior means are obtained from the $\mathbf{p}_{opt}$ solution



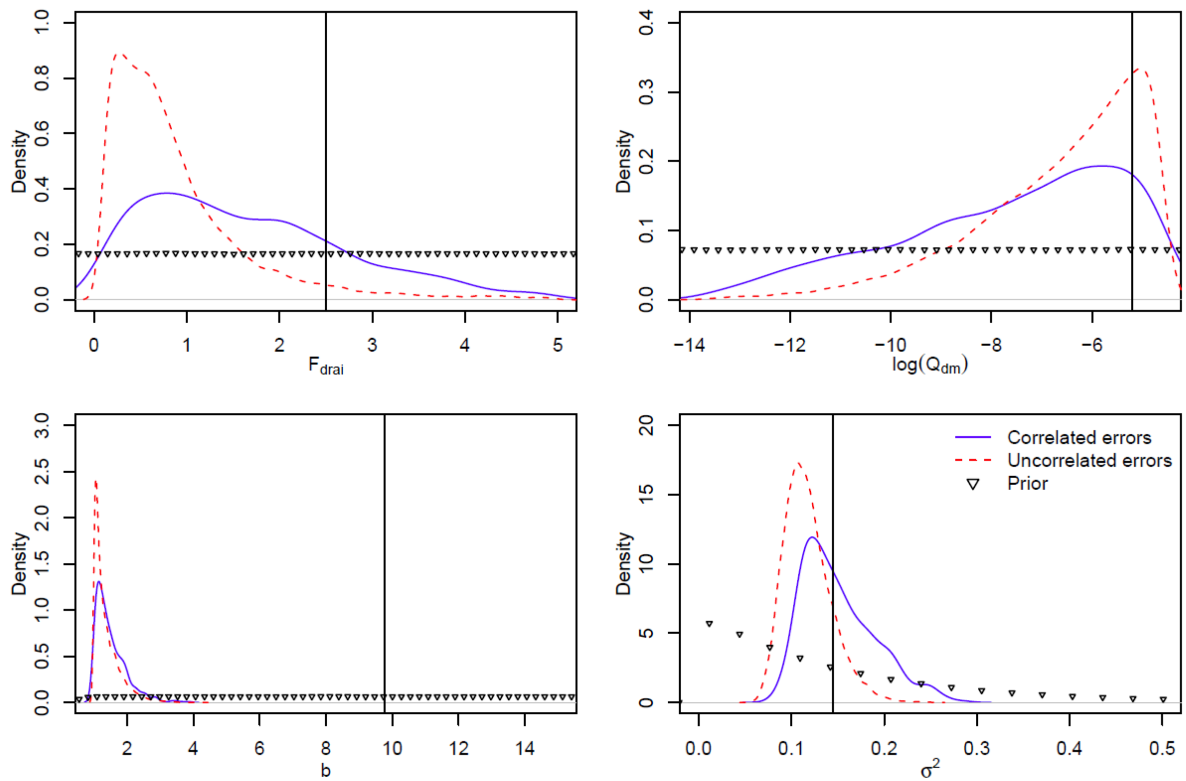
Solving the Bayesian calibration problem

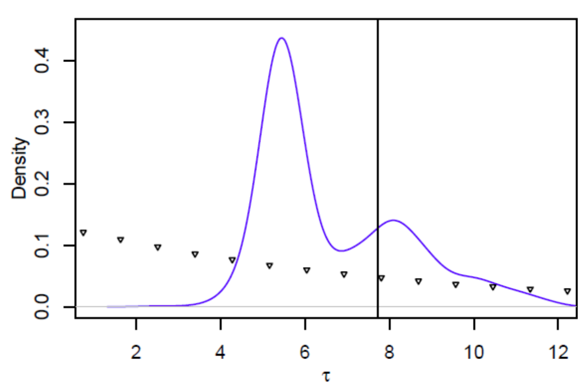
- Conventional formulation using Bayes' theorem and independent priors for all parameters
  - 4 parameter inversion when using uncorrelated errors; 5 with correlated ones
- 10^5 model invocations needed for converged parameter posterior distributions
  - Convergence checked with Raftery-Lewis statistic
  - Calibration quality checked using (1) posterior predictive tests and (2) verification rank histogram
- Check the following
  - How far off are the nominal/default parameter values?
  - Is $\mathbf{p}_{opt}$ contained inside the posterior PDFs?
  - Which model gives a smaller structural error – correlated or uncorrelated?
  - What is the time-scale of correlated structural error?

Sandia National Laboratories

Posterior distributions

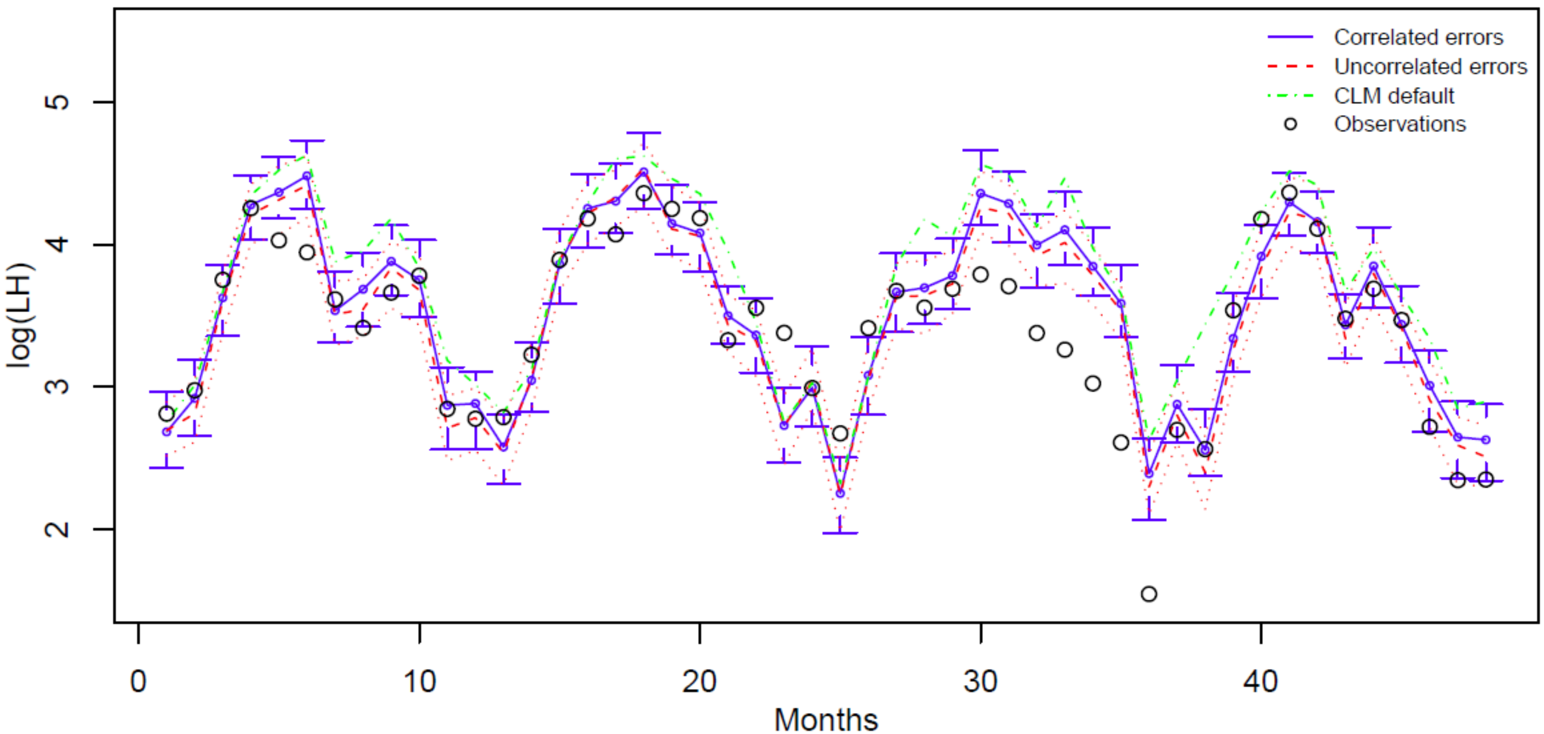
- Default parameters very suboptimal
- $\mathbf{p}_{opt}$ in the PDF's support





- Uncorrelated structural error better
- Correlation time-scale ~5 months

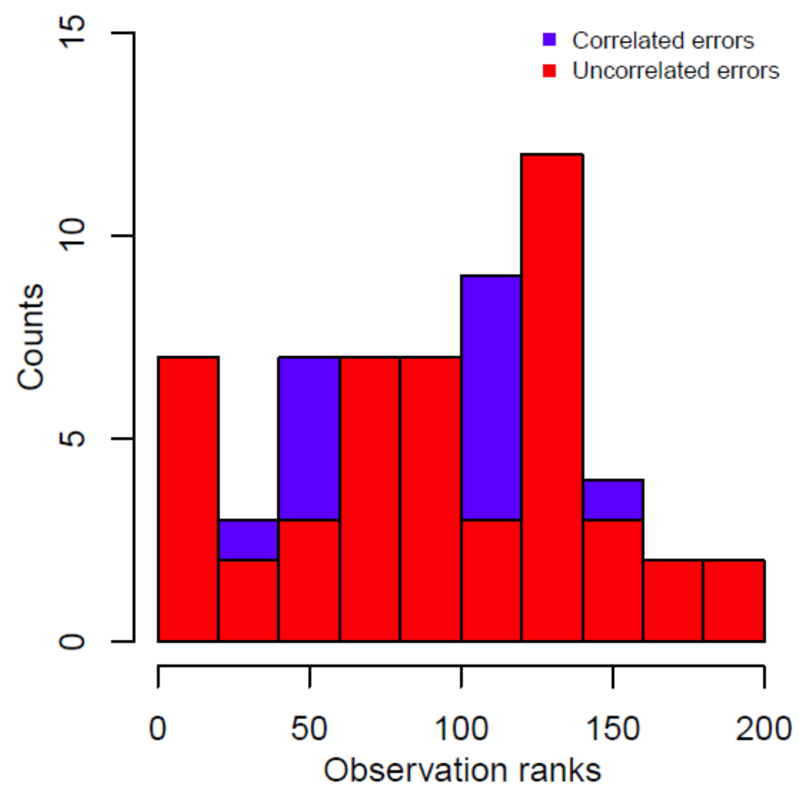
Posterior predictive test



- Most observations are captured

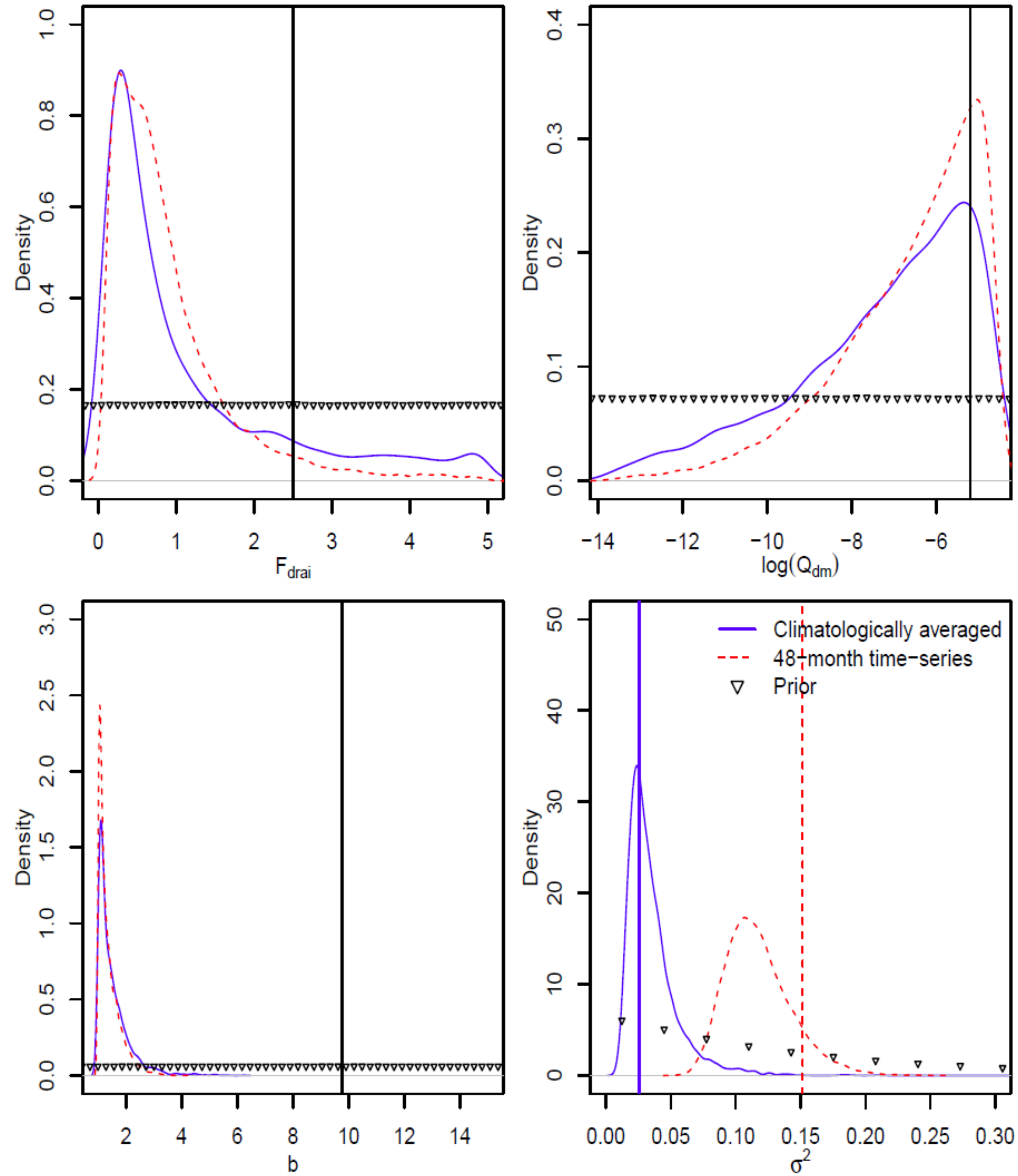
Quality of the calibration

- Ideally the distribution of observation ranks should be uniform
- If bunched in the middle, over-dispersive calibration
- In our case, ranks bunched towards the higher side
  - Under-dispersive calibration
- Choice of structural error model is immaterial



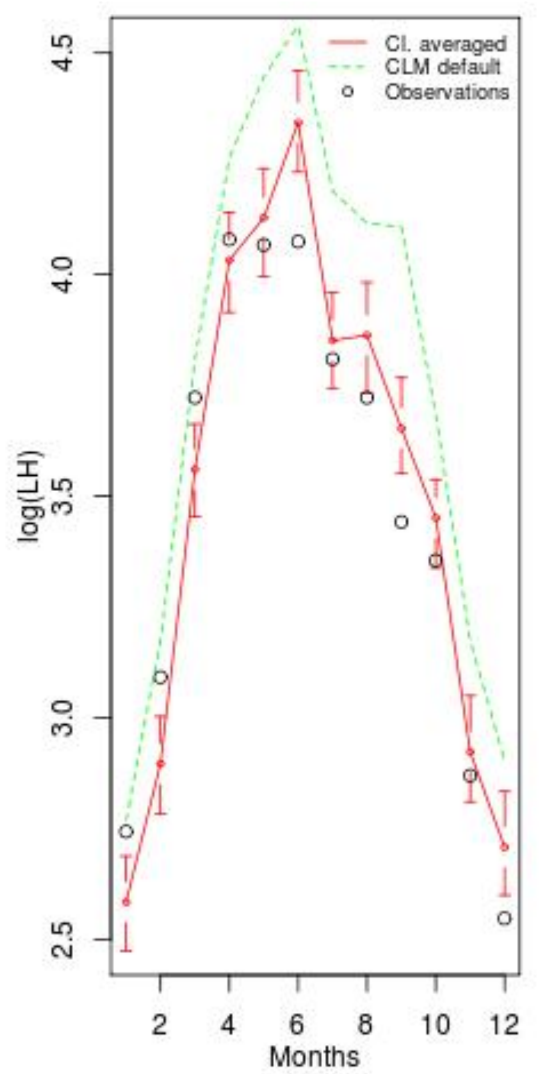
Effect of climatological averaging

- Climatological averaging results in a 12-month time-series
- Not much change in parameter PDFs
  - CLM is probably not meant to capture year-to-year variability



Posterior predictive test

- Big improvement over the default parameter values
- Error bars capture most of the observations



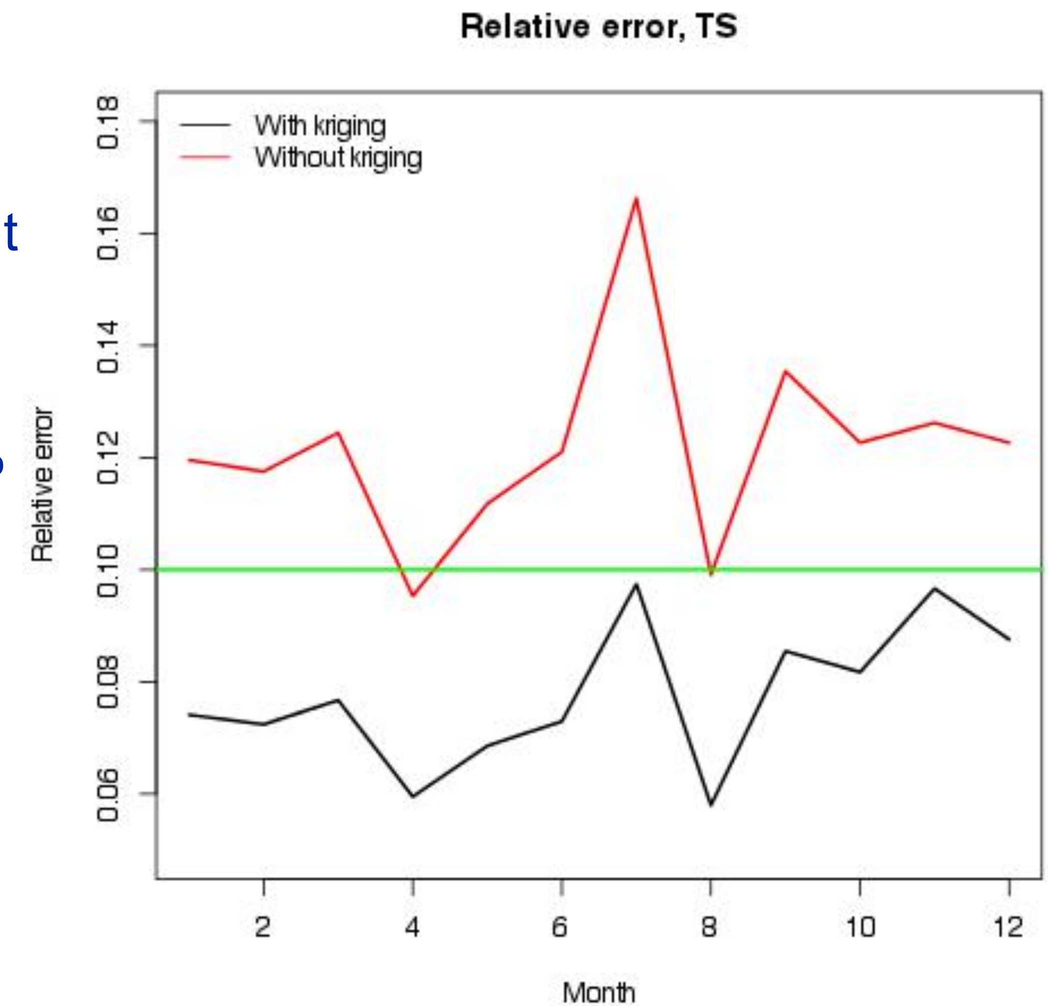
Surrogate modeling and calibration

US-MOZ

Sandia National Laboratories

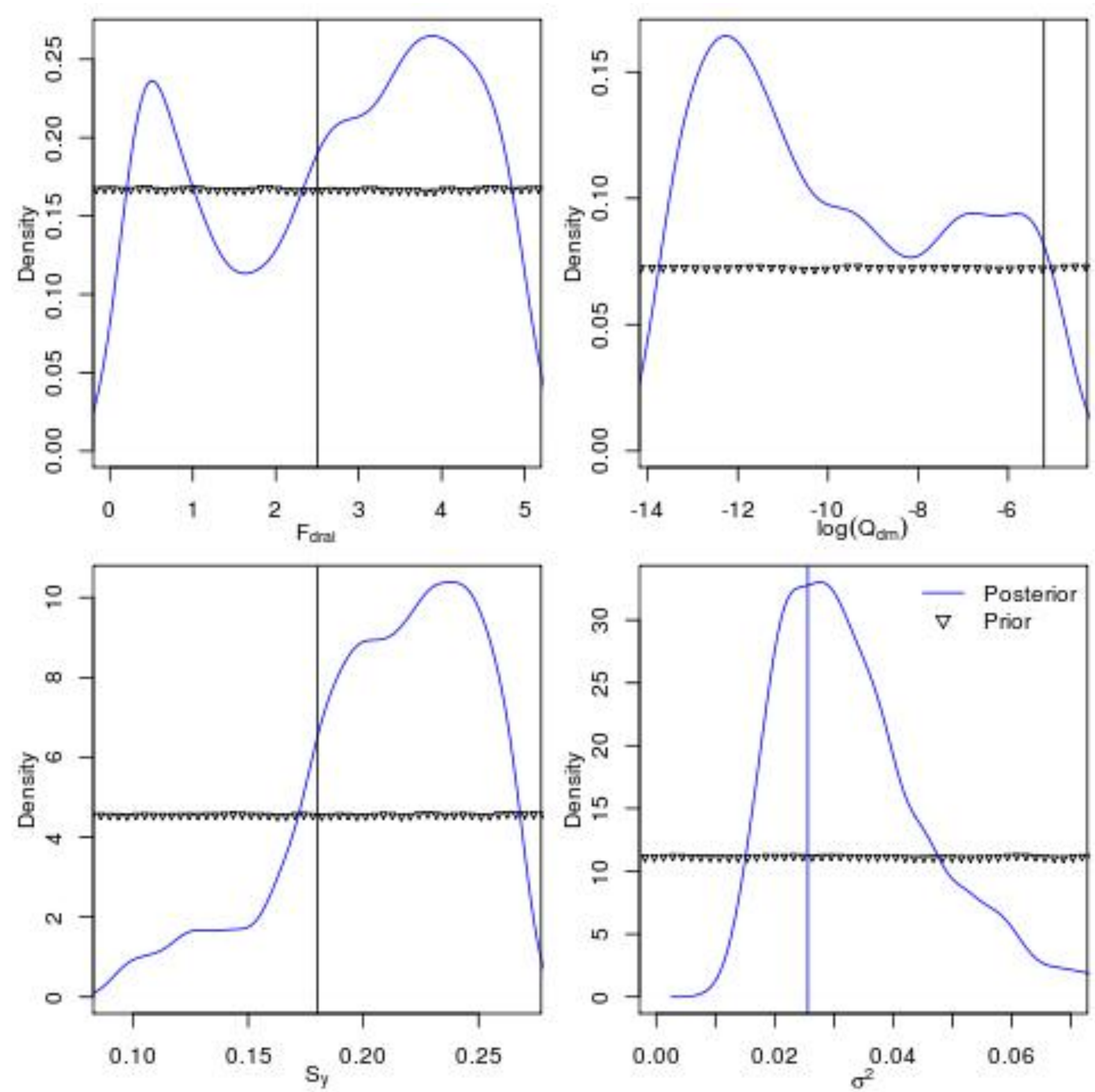
Surrogate models

- Parameters: $\{F_{drai}, \log(Q_{dm}), S_y\}$
- Could not make surrogate models without climatological averaging
- Quadratic models after averaging had 15%-20% errors
- Modeled defect as a 3D mGaussian field
  - Exponential variogram
- Regression kriging model

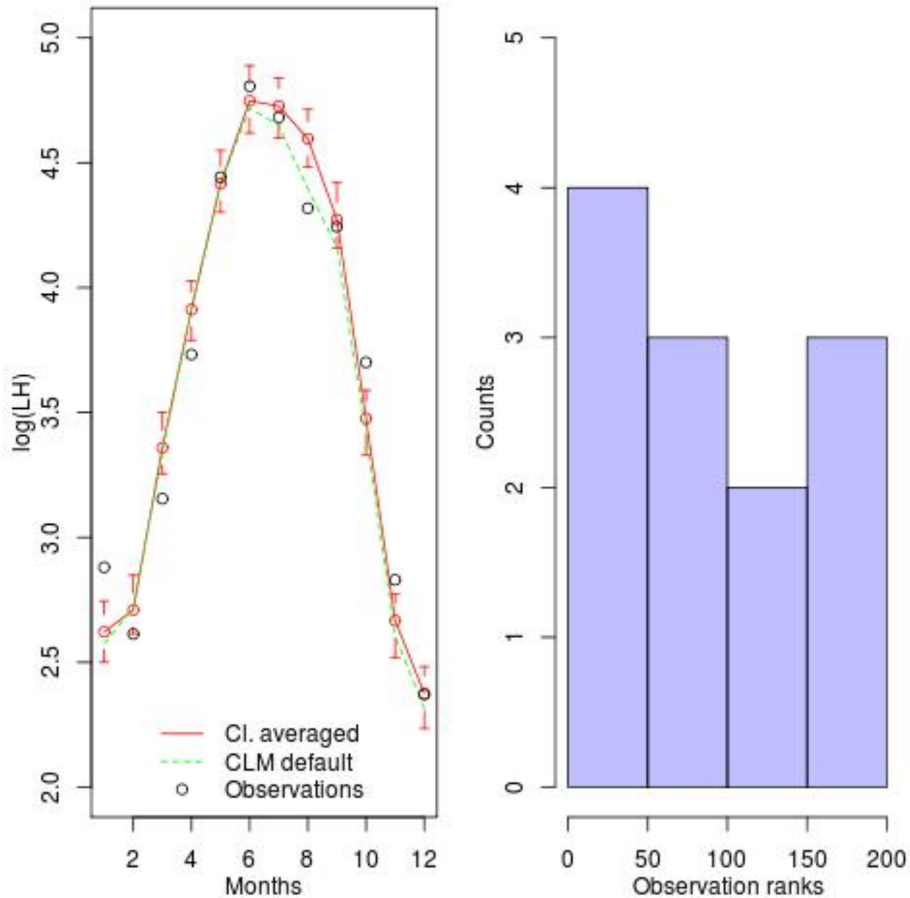


Parameter estimation

- $\mathbf{p}_{opt}$ = {2.6, log(4.43 × 10−3), 0.2 }
- Default & optimal parameter values not near PDF peaks
- Model-data mismatch accurately estimated by both deterministic & Bayesian approaches



Posterior predictive tests



- Calibration captures observations
- Not much improvement over default parameter values

Conclusions

- It is possible to create fast-running surrogates of CLM
- In such cases, we can
  - Perform Bayesian calibration of parameters and obtain their PDFs
  - Estimate the structural error, under various error model
- The structural error tells us
  - The accuracy of CLM, post calibration
  - Some characteristics of the physical processes that cause the structural error
- The most difficult aspect of the study is proposing a model form for the surrogate models
  - Requires testing & improvement on a month-by-month basis
  - No guidance on what model form might work for a given month
- *Details: J. Ray, Z. Hou, M. Huang and L. Swiler, "Bayesian calibration of the Community Land Model using surrogate models" Sandia Technical Report, SAND2014-0867*

Sandia National Laboratories

BACKGROUND

Current verification/validation status of CLM

- Sensitivity analysis – lots of work
  - Total ~80 parameters
  - Most sensitive parameter depend on (1) observable and (2) location
  - For LH, hydrological process are most important; captured by 3 parameters. Different parameters for different sites
- Bayesian calibration of CLM
  - 1 published paper; calibrated 10 parameters using LH and runoff measurements
  - MCMC using CLM as-is; expected values of parameters were obtained (converged values)
- Yet to be done
  - Construction of PDFs, not expected values, of CLM parameters from observations
  - Structural error estimation

Usefulness of calibration with CLM

- Calibrated parameters improve predictive skill
- Doesn't matter whether we use surrogates or CLM for the prediction