Quantification of structural uncertainty in a land surface model

Z HOU, M HUANG, PNNL; J RAY, L SWILER, SNL

Pacific Northwest National Laboratory
Sandia National Laboratory
Mar 18 2015
Motivations
- To identify and quantify model structural uncertainty in the Community Land Model

Approaches
- Efficient sampling–based sensitivity analysis
- Classification of complex climate system
- Time-frequency analysis of ensemble simulation errors
- Separation of model parametric uncertainty and structural uncertainty
  - MCMC-Bayesian using numerical forward models and/or surrogates
  - Time-frequency analysis of simulation errors from posterior samples
Motivations

- Climate system: multi-phase, multi-component, multiple biogeophysical/chemical processes
- Numerous model and coupling parameters; formidable high-dimensional parameter spaces
- Uncertain parameter values (parametric uncertainty)
- Model structural uncertainty (makes parameter inversion questionable)
Assumptions

Sources of uncertainty

- Model uncertainty
  - Simplifications, structural model formulations/structures, extrapolations, resolution, model initial/boundary conditions

- Parameter uncertainty
  - Non-measurable, measurement errors, non-uniqueness, inaccurate calibration, misclassification due to under-sampling...

- Data uncertainty
  - Instrumental errors, consistency, gaps, resolution, scaling

- Natural uncertainty/variability/heterogeneity
  - Intrinsic quantities vary over time, over space, or across individuals in a population
  - Physical processes/mechanisms/features vary over space, time, and individuals

Assumptions

- No systematic data measurement errors
- Model prediction errors are mainly due to parametric and model structural uncertainty
- The complex climate system can be divided into simpler subsets/groups, with each has common model structural errors
Step 1: Entropy concept and efficient sampling to fully represent parametric input uncertainty

### Prior information

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{\text{max}}$</td>
<td>Max fractional saturated area, from DEM</td>
</tr>
<tr>
<td>$C_s$</td>
<td>Shape parameter of the topographic index distribution</td>
</tr>
<tr>
<td>$f_{\text{over}}$</td>
<td>Decay factor (m$^3$) for fsat</td>
</tr>
<tr>
<td>$b$</td>
<td>Clapp and Homberger exponent</td>
</tr>
<tr>
<td>$K_r$</td>
<td>Hydraulic conductivity (mm s$^{-1}$)</td>
</tr>
<tr>
<td>$\theta_s$</td>
<td>Porosity</td>
</tr>
<tr>
<td>$\psi_s$</td>
<td>Saturated soil matrix potential (mm)</td>
</tr>
<tr>
<td>$f_{\text{dual}}$</td>
<td>Decay factor (m$^3$) drainage</td>
</tr>
<tr>
<td>$q_{\text{dual}}$</td>
<td>Max drainage (kg m$^{-2}$ s$^{-1}$)</td>
</tr>
<tr>
<td>$S_y$</td>
<td>Average specific yield</td>
</tr>
</tbody>
</table>

### Entropy concept

$$f(x) = \frac{\sqrt{\gamma}}{\pi} \exp \left[ -\gamma \left( x + \frac{\beta}{2\gamma} \right)^2 \right] \Phi \left[ \sqrt{2\gamma \left( U + \frac{\beta}{2\gamma} \right)} \right] - \Phi \left[ \sqrt{2\gamma \left( L + \frac{\beta}{2\gamma} \right)} \right]$$

$L \leq x \leq U$

### Prior pdfs

Efficient sampling
Step 2: Classify the complex system into subsets/groups

Parameter identifiability

Clusters of US MOPEX basins

Expectation-maximization clustering
Step 3: Time-frequency decomposition of prior ensemble simulation errors

Wavelet analysis of simulation errors

Model simulation errors
Step 3: Time-frequency decomposition of prior ensemble simulation errors
Step 4: Reduction of parametric uncertainty via MCMC-Bayesian inversion

- MCMC-Bayesian inversion
- Surrogate development
- Parameter subspace selection
- Data refinement/classification

(Ray et al. 2015 and talks in sessions MS159/MS162)
Step 5: Time-frequency decomposition of posterior ensemble simulation errors

Power spectrum of simulation errors before parameter inversion

Power spectrum of simulation errors after parameter inversion
It is possible to identify model structural errors by quantifying input parameter uncertainty and fully exploring the input parameter space.

(Ensemble) model simulation errors provide information about the processes (and/or parameters) with major contributions to the errors.

Assuming no systematic errors in the conceptual models and observational data, the model structural errors can possibly be separated after parametric uncertainty is reduced.

The remaining errors would provide guidance on further model improvement, e.g., by modifying the physical models or parameterizations that numerically affect the errors at the major spatial-temporal scales.