Calibrating and Re-Calibrating a Global Vegetation Model

Thomas G. Dietterich
Eric Chown
MAPSS (Neilson, 1995)

Predictions
- LAI in each pixel ($\text{LAI}_{\text{tree}}$, $\text{LAI}_{\text{grass}}$, $\text{LAI}_{\text{shrub}}$)
- Classification into one of 74 possible biomes
  - tall grass prairie, desert, conifer forest, etc.
MAPSS (2)

Optimization model:
- find LAI values that achieve “water balance” and “light balance”

Water Balance
- \((\text{soil water in} - \text{soil water out}) = 0\) over 12 months
- soil water always \(\geq 0\)

Light Balance
- amount of light reaching forest floor matches grass LAI
Hydrology Model

Rainfall-Snowfall
- canopy interception and through-fall
- snow accumulation and melt
- separate modeling of saturated and unsaturated soil water flow
- separate modeling of deep, medium, and shallow soil water pools
- transpiration from shallow soil for grasses and shrubs, from shallow and medium for trees

Each 10x10km cell is independent (no lateral flows)
Calibration

- MAPSS is an aggregate model
- Transpiration and soil water flow equations chosen empirically
  - $\alpha e^{\beta X}$
  - must set two parameters in each equation
- Criteria
  - correctly predict boundaries of major biomes
  - correctly predict seasonality of water flows
  - correctly predict observed LAI
    - 22 quantitative sites; overall qualitative behavior
Manual Calibration

- Calibrate grass transpiration parameters and unsaturated deep and middle water parameters
  - use data from praries, where trees and shrubs are absent and there is no saturated water flow
- Calibrate tree transpiration parameters
  - forests with unsaturated soil only
  - separately for different climate zones
- Calibrate shrub transpiration parameters
  - shrub savannah with unsaturated soil
- Calibrate top layer saturated water flow
  - shrublands where middle and deep water flow are “lost”
- Calibrate deep saturated water flow
  - grasslands where middle flow can be ignored
- Calibrate middle saturated water flow
  - grasslands where all flows occur
Automated Re-Calibration

Goal: Determine how stable the model parameterization is

Method:

– Use calibrated model (known parameter values) to generate predicted LAI values over USA

– Apply optimization algorithm to see if we can recover these parameter values
Automated Calibration

Define error measure

\[ J(\Theta) = \left( \widehat{LAI}_{tree} - LAI_{tree} \right)^2 + \left( \widehat{LAI}_{grass} - LAI_{grass} \right)^2 + \left( \widehat{LAI}_{shrub} - LAI_{shrub} \right)^2 + \left( \widehat{RUNOFF} - RUNOFF \right)^2 \]

Search for \( \Theta \) to minimize \( J \)
Global Optimization Algorithms

- Non-gradient search (Powell’s method)
- Gradient search (conjugate gradient)
- Simulated Annealing

- All FAIL on this problem!
  - nonlinearlities and complex interactions
Automated Model Decomposition

- Identify parameter subsets that can be calibrated independently (sequentially)
- Identify sites that we are confident correspond to those parameter subsets
- Apply simulated annealing to parameter subsets
Method

- Automated program analysis to identify paths through the simulation that only involve small numbers of parameters
- Empirical method of identifying data points that belong to a path (with high probability)
- Optimization of parameter subsets
## Transpiration Parameter Results

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Grass</th>
<th>Tree1</th>
<th>Tree2</th>
<th>Tree3</th>
<th>Shrubs</th>
<th>SlopeG</th>
<th>SlopeT</th>
<th>SlopeS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.265</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.792</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4.249</td>
<td>3.748</td>
<td></td>
<td></td>
<td></td>
<td>1.417</td>
<td>1.001</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4.259</td>
<td></td>
<td>3.502</td>
<td></td>
<td></td>
<td>4.999</td>
<td>1.003</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4.278</td>
<td></td>
<td></td>
<td>2.750</td>
<td></td>
<td>3.496</td>
<td>1.032</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.252</td>
<td></td>
<td></td>
<td></td>
<td>9.254</td>
<td>2.248</td>
<td></td>
<td>3.002</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.246</td>
<td></td>
<td></td>
<td>3.002</td>
</tr>
<tr>
<td>7</td>
<td>4.249</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.998</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>4.251</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.072</td>
<td></td>
</tr>
<tr>
<td>Target</td>
<td>4.250</td>
<td>3.750</td>
<td>3.500</td>
<td>2.750</td>
<td>9.275</td>
<td>2.000</td>
<td>1.000</td>
<td>3.000</td>
</tr>
<tr>
<td>Range</td>
<td>2 – 8.5</td>
<td>1.5 – 5</td>
<td>1.5 – 5</td>
<td>1.5 – 5</td>
<td>1.2 – 4</td>
<td>4 – 15</td>
<td>0 – 5</td>
<td>0 – 5</td>
</tr>
</tbody>
</table>
## Soil Water Flow Results

<table>
<thead>
<tr>
<th>Iter</th>
<th>Deep1</th>
<th>Deep2</th>
<th>Mid1</th>
<th>Mid2</th>
<th>Top1</th>
<th>Top2</th>
<th>Deep1</th>
<th>Deep2</th>
<th>Mid1</th>
<th>Mid2</th>
<th>Top1</th>
<th>Top2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.786</td>
<td>11.476</td>
<td>0.286</td>
<td>2.744</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.704</td>
<td>11.607</td>
<td>0.125</td>
<td>2.337</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.194</td>
<td>9.971</td>
<td>0.491</td>
<td>2.994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.443</td>
<td>10.934</td>
<td>0.597</td>
<td>2.939</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.371</td>
<td>10.293</td>
<td>0.640</td>
<td>5.051</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.443</td>
<td>10.933</td>
<td>0.604</td>
<td>2.944</td>
<td>0.834</td>
<td>2.530</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.274</td>
<td>9.754</td>
<td>0.603</td>
<td>3.644</td>
<td>0.849</td>
<td>11.447</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.201</td>
<td>9.982</td>
<td>0.504</td>
<td>3.003</td>
<td>0.857</td>
<td>10.128</td>
<td>0.816</td>
<td>3.087</td>
<td>0.613</td>
<td>0.719</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target</td>
<td>0.200</td>
<td>10.000</td>
<td>0.500</td>
<td>3.000</td>
<td>0.800</td>
<td>2.500</td>
<td>0.800</td>
<td>10.000</td>
<td>0.800</td>
<td>3.000</td>
<td>0.500</td>
<td>1.000</td>
</tr>
<tr>
<td>Range</td>
<td>0 – 1</td>
<td>1 – 20</td>
<td>0 – 1</td>
<td>1 – 7.5</td>
<td>0 – 1</td>
<td>1 – 7.5</td>
<td>0 – 1</td>
<td>1 – 20</td>
<td>0 – 1</td>
<td>1 – 5</td>
<td>0 – 1</td>
<td>0.5 – 3</td>
</tr>
</tbody>
</table>
Lessons Learned

- Data is too sparse to support fully-automated calibration
- Complex models cannot be calibrated simply by wrapping a clever optimization algorithm around the system (black box optimization)
- Automated re-calibration is possible, but requires a divide-and-conquer strategy
- Associating data with paths through a complex model can be automated
- Calibration of model subcomponents is possible, but requires extensive hand-tweaking of optimization parameters
- MAPSS calibration is stable except for middle layer saturated flow parameters, which are under-constrained.