Methods for spatial upscaling and risk analysis of process-based models

Experiences from projects GREENHOUSE and Carbo-Extreme

Marcel van Oijen, David Cameron, Peter Levy (CEH-Edinburgh), and colleagues in GREENHOUSE and Carbo-Extreme

Reading, 9 Apr 2015
Spatial upscaling methods used for process-based models

Individual grid cells ("regions") are large → What is the model output uncertainty for each cell?

UPSCALING METHOD

Input aggregation

\[ \overline{y} = f(\overline{x}) \]

Point selection

\[ \overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \]

Interpolation (e.g. geostatistics)

\[
p(\overline{y}|\{y_i\}) = \int p(\overline{y}|\theta, \{y_i\}) p(\theta|\{y_i\}) d\theta
\]

\[ \Theta = \text{Spatial hyperparameters (variance, correlation length)} \]
Example: region 5 x 5 km

Model:

PAR (mumol m$^{-2}$ s$^{-1}$)

GPP (mumol m$^{-2}$ s$^{-1}$)
UPSCALING METHOD 1: Input Aggregation

Biased estimates of regional mean
UPSCALING METHOD 2: Point Selection

GPP from selecting random points (mumol m-2 s-1)
Point selection 1

Upscaling by point-selection

# points
UPSCALING METHOD 3: Bayesian Kriging

Uncertainty maps
EXAMPLE 1: Comparison of methods

Biased method

Unbiased methods
Example region: 128 x 64 km
Model BASECO

Atmospheric drivers
- N-deposition
- CO₂
- Radiation
- Temperature
- Rain
- Humidity
- Wind speed

Parameters & initial constants
- vegetation
- soil

Input → Model → Output
Some inputs and outputs of BASECO
EXAMPLE 2: Comparison of methods

**N$_2$O**
(kg N ha$^{-1}$ y$^{-1}$)

**NEE**
(kg N ha$^{-1}$ y$^{-1}$)

**GPP**
(kg N ha$^{-1}$ y$^{-1}$)
Spatial heterogeneity: Gaussian Random Fields (GRFs)

Variance

Spatial correl. length (clustering)
A.

B. Smaller spatial correlation length (less clustering)

C. Greater variance
Current work: upscaling for multiple land-use types
• **Upscaling errors** depend on:
  • sampling intensity and spatial distribution
  • output variable
  • variance & autocorrelation length

• Input Aggregation gives **biased** estimates of the regional mean; the other methods are **unbiased**

• **Uncertainty estimation** is part of (geostatistical) Interpolation; the other methods require data from representative regions
Risk analysis in project Carbo-Extreme


How does the analysis work?

\[ \text{RISK} = \text{Average loss} \]

\[ \text{RISK} = \text{VULNERABILITY} \times \text{P(Hazardous)} \]

\[ \text{P(Hazardous)} = \text{frequency } env < \text{thr} \]
Risk analysis carried out in Carbo-Extreme

- **Time periods:** 1971-2000, 2071-2100
- **sys:** NPP, NEP, $R_h$, SWC, ET
- **env:** Drought (*)

**Models:**

<table>
<thead>
<tr>
<th>Group</th>
<th>Model</th>
<th>Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIK</td>
<td>LPJ</td>
<td>generic</td>
</tr>
<tr>
<td>LCSE</td>
<td>ORCHIDEE</td>
<td>generic</td>
</tr>
<tr>
<td>MPI</td>
<td>JSBACH</td>
<td>generic</td>
</tr>
<tr>
<td>CEH</td>
<td>BASFOR</td>
<td>forest</td>
</tr>
<tr>
<td>IIASA</td>
<td>EPIC</td>
<td>crops</td>
</tr>
<tr>
<td>INRA</td>
<td>PASIM</td>
<td>grassland</td>
</tr>
</tbody>
</table>

(*) Threshold: Standardized Precipitation Evapotranspiration Index (SPEI) < -1
Example: Risk analysis for NPP

NPP in good years

→ VULNERABILITY
   (= Difference NPP good-bad)

→ RISK
   (= Prob.*VULNERABILITY)

NPP in bad years

→ Probability of bad years
Risk analysis NPP (model PIK) 1971-2000

A: $E[\text{NPP} \mid \text{SPEI} > -1]$ 1971-2000
B: $E[\text{NPP} \mid \text{SPEI} < -1]$ 1971-2000
C: Vulnerability $[\text{NPP} \mid \text{SPEI} < -1]$ 1971-2000
D: $P[\text{SPEI} < -1]$ 1971-2000
E: Risk $[\text{NPP} \mid \text{SPEI} < -1]$ 1971-2000
Risk analysis NPP (model PIK) per latitudinal band

- Latitude > 55° N
- 45° < lat. < 55° N
- Latitude < 45° N

1971-2000-2100
Risk analysis NPP: model comparison

- PIK (generic)
- LCSE (generic)
- MPI (generic)
- IIASA (wheat)
- CEH (forest)
• Definitions of risk (R), vulnerability (V) and probability of hazard (P(H)) allow decomposition: 
  \[ R = V \times P(H) \]

• Analysis method is applicable to data as well as to model outputs

• For drought, both V and R expected to increase in the south, but large model (vegetation) differences