Combining physical and statistical models to predict environmental processes.

Jim Zidek

U British Columbia
Acknowledgements

- Prasad Kasibhatla, Duke
- Douw Steyn, UBC
- Co-authors:
  - Nhu Le, BC Cancer Agency
  - Zhong Liu, Capital One
Outline

- PART I: Some relevant UBC research
- PART II: Phystat modelling - fundamentals
- PART III: Phystat modeling - approaches
- Conclusions
PART I: UBC RESEARCH
Current research involving natural resources.
NICDS - Agri & Agri - Food Canada

Soil - water - climate change - food - biofuel
NICDS project with AAFC partnership 2007 -

- One year’ only of NICDS funding - failure of NSERC renewal. Success story. Work continues.

- AAFC has provided:
  - scientific collaboration
  - positions: 2 PDFs; 1 year full RA; MSc coop
  - data
  - interesting projects
  - meeting opportunities
  - cyber course instruction
Projects completed:

- Markov models for binary climate processes (PhD thesis)
- Phenological phenomena models, e.g. bloom dates for grape vines (MSc Thesis)
- Crop yield forecasting models based on soil - moisture characteristics (Current PhD thesis).
Conventional regression residuals: yields on soil moisture by agrodistrict. Fails to borrow strength by exploiting spatial modeling techniques.
Future work:

- Future crop yields based on downscaled climate model estimates
- Web portals
- Design of micro sensor monitoring networks - soil conditions.
FPInnovations

Forests - climate change Current NSERC - CRD - FPInnovations research grant. 5 year collaborative research program - sited at UBC in partnership with SFU - concerns strength lumber. Current projects:

- design of sampling programs - cross sectional and longitudinal - catastrophic change like mountain pine beetle - long term trends due to climate change.
- property relationships - e.g. cracks vs bending strength
- duration of load = accelerated testing
  - integrates deterministic engineering models with data
FPInnovations

Forests - climate change
Possible opportunities for cross Canada collaboration as FPI has labs in both the East and West.
PART II: PHYSTAT MODELLING - FUNDAMENTALS
Origins

Need to model environmental space-time fields over large space-time domains that challenge physical and statistical modelers.
Environmental Space-Time Fields, $X$

$x$ massively multivariate: over $\text{time} \times \text{space} \times \text{species}$, often discretized:

- time may mean, hour, day, month, etc
- space can refer to a point with a latitude and longitude or a region eg a county
- species:
  - chemical species, gases, aerosols, etc
  - generally, any vector of dependent variables eg 24 hourly values within day

What’s a Model?

“an abstract, analogue representation of the prototype whose behavior is being studied” (Steyn & Galmarini 2003)
Simulation Model Taxonomy

**Analytic Models:**
- variables in tractable math equations represent measurable attributes of the real thing

**Physical Scale Models**
- physical behavior of their measurable properties analogous to that of the real thing

**Numerical Models**
- variables obtained by numerical solution thought to be analogous to measurable attributes of the real thing. _Model outputs called “simulated data”._
Models: The problem of meaning

Controversy! The Oreskes Paper

- highly influential
  - says physical models cannot be shown to represent reality - validation meaningless/pointless
  - still cited over 40 times per yr
  - used to justify not validating!
Models: The problem of meaning

Controversy! The Oreskes Paper

- dismisses common assessment practices
  - verification
  - validation
  - verifying numerical solutions
  - calibration
  - confirmation
Confirmation: match between simulated and real data implies verification (truth) - logical fallacy called “affirming the consequence”

EXAMPLE: Hypothesis H: “It is raining.” Model: “If H, I will stay home and revise the paper.” You find me at home and therefore conclude it is raining because empirical data matches predicted outcome under the hypothesis of model!

Failing to predict the data ⇒ bad model. Success ☞ good model! Moreover, numerous models could predict the same observations equally well!
Summary:

“The primary purpose of models in heuristic...useful for guiding further study but not susceptible to proof... [Any model is] a work of fiction (OUB citing philosopher Nancy Cartwright). ... A model, like a novel may resonate with nature, but is not the ‘real thing’.”
Models: The problem of scales

Combining simulated & Real Data: Does it make sense?

Example:

\[
\frac{1 + 1}{2} = 0.5
\]

Seems correct. But it's actually nonsensical.
Models: The problem of scales

Combining simulated & Real Data:
Does it make sense?

Example:

\[
\frac{(1 + 1)}{2} = 0.5
\]

Seems correct. But its actually nonsensical.

\[
\frac{(1 \text{ cm} + 1 \text{ apple})}{2} = 0.5
\]

Nonsensical! Simulated data also on different scales than real data.
Model Dynamic Scales

Steyn & Galmarini 2003 demonstrate the problem.

Continuous real data monitors on a space - time scale of just $1 \, m^2 \times \text{few minutes}$ in lower left hand corner!!
PART III: PHYSTAT MODELLING - APPROACHES
Motivating Example: Calibration

Air Pollution → Disease/Death

Need to regulate → Need to monitor

Need to relate ambient to personal exposures

Need dose response models

Impact estimates/control!!!
Ozone example

Pollution:

Sources:

US Ozone

NA Anthro

Not NA Anthro
Ozone example

Policy Related Background (PRB)

Not NA Anthro

Foreign Anthro

Natural Sources

Lightning

Wildfires

Biogenic Emissions
Estimating the PRB

- **Not measurable:**
  - urban pollution spreads to rural areas
  - few pristine sites available
  - representative of contaminated areas???

- **Alternative:** infer from deterministic chemical transport models (CTMs)
  - **GEOS-CHEM** used for ozone
  - **MAQSIP** similar (see below)

- **Calibration needed:** to represent “ground truth” (measurements)
Calibrating CTMs

Fundamental issue: different scales.

CTMs: meso-scale models
Field measurements: microscale
Bayesian melding approach

Bayesian melding model (BM) (Fuentes & Raftery, Biometrics, 2005):

\[
\hat{Z}(s) = Z(s) + e(s)
\]

\[
\tilde{Z}(s) = a(s) + b(s)Z(s) + \delta(s)
\]

\[
\tilde{Z}(B) = \int_{B} \tilde{Z}(s)ds
\]

\[
\tilde{Z}(B) \approx \frac{1}{L} \sum_{j=1}^{L} \tilde{Z}(s_{j,B}).
\]

- \( \{s_{j,B}\} \): sampling points within \( B \);
- \( \hat{Z}(s) \): measurements;
- \( \tilde{Z}(B) \): model outputs; \( Z(s) \): “true” process.
Model calibration

Calibration formula:

Recalibrated $\tilde{Z}(B) = \left( \tilde{Z}(B) - \frac{1}{|B|} \int_B a(s) ds \right) / b.$

Each iteration of the MCMC generates a recalibrated value & overall, empirical marginal distribution for it.
BM calibration assessment

How well does recalibrated model outputs predict measurements?

- **Uses 10AM - 5PM averages:** measurements & MAQSIP model outputs (Kasibhatla & Chameides 2000; Fiore et al. 2003, 2004)

- **Model outputs:** from 78 grid cells.
- **Measurements:** from 48 stations.
- **Validation data:** measurements from remaining 30 stations (all collocated with grid cells by choice) to be predicted.
BM calibration assessment

How well does recalibrated model outputs predict measurements?
Define: root mean square prediction error (RMSPE):

\[ \sqrt{\frac{1}{n} \sum_{i}^{n} (O_i - \hat{O}_i)^2}, \]

\( O_i \) = measurement at location \( i \); \( \hat{O}_i \) = prediction at location \( i \).

Results:

<table>
<thead>
<tr>
<th></th>
<th>melding</th>
<th>Kriging 1</th>
<th>Kriging 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>13.37*</td>
<td>14.24</td>
<td>14.79</td>
</tr>
</tbody>
</table>

Average RMSPE for 30 days. **Kriging1**: Using measurements. **Kriging2**: Using model outputs.
Univariate STReg approach

Univariate spatial - temporal regression

- Extends Guillias et al. (2006)
- Ignore misaligned measurements - to - model supports.
- Relate measurements \( \{O(t)\} \) to mod outputs \( \{M(t)\} \) at each monitoring site:

\[
O(t) = c + aM(t) + N_t, \quad t = 1, 2, \cdots, T
\]
\[
N_t = \rho N_{t-1} + \epsilon_t
\]
\[
e_t = \sum_{i=1}^{p} \gamma_i Z_i(t) + \epsilon_t
\]
\[
\epsilon_t \sim N(0, \sigma^2_\epsilon) \ i.i.d
\]
Univariate STReg approach

Univariate spatial - temporal regression

- Make **coefficients** site dependent & have joint Gaussian field: $a$, $c$, $\gamma_i$ & residuals $\epsilon$.
- **Temporal process model**: AR($p$) with spatially correlated coefficients & residuals.
- Gives spatial predictions & temporal forecasts.
Model calibration: Uni STReg

Calibration formula:

\[ \tilde{Z}(B) = a + c\tilde{Z}(B). \]

Incorporated into MCMC runs.
Univariate STReg assessment

- **As before:**
  - 78 stations collocated with 78 model grid cells
  - use all 24 hours, not just 8-hour averages.

- **Fit model:** use first 240 hours, measurements at 15 monitoring sites for this.

- **Prediction & forecasting:** use all model outputs, 78 stations & 480 hours.

- **For the 15 stations:** forecast future in $2^{nd}$ block of 240 hours.

- **Spatial predictions:** 1$^{st}$ 240 hour measurements for remaining 63 sites.
Uni STReg assessment

Overall Uni STReg calibration beats Kriging, Melding (small number of stations - 15 to predict 63; only spatial models only).

- **Overall accuracy**: Uni STReg does better, spatial prediction & forecasting.
- **Coverage probabilities**: 90% credible intervals for forecasts = 90%; for predictions 95% (too big!).

**RESULTS:**

<table>
<thead>
<tr>
<th></th>
<th>Uni STReg</th>
<th>Model outputs alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSPE</td>
<td>14.43</td>
<td>16.50</td>
</tr>
<tr>
<td>RMSFE</td>
<td>15.57</td>
<td>18.57</td>
</tr>
</tbody>
</table>
Multivariate STReg

**Extends Uni STReg:** Can borrow strength even in univariate case so improves on Uni STReg.

\[
O_{s,t} = \beta_s M_{s,t} + N_{s,t}
\]
\[
N_{s,t} = \rho N_{s,t-1} + \gamma_s Z_{s,t} + \epsilon_{s,t}
\]
\[
\epsilon_{s,t} \sim \text{MVN}(0, \Sigma_\epsilon)\]

- **O_{s,t}**: \(q \times 1\) measurements vector, \(q\) pollutants.
- **M_{s,t}**: \((p + 1) \times 1\), vector of intercept terms, & \(p\) model outputs, for the \(q\) pollutants.
- \(p = q\) not necessary.
Multivariate STReg

- Make $\beta_s$ & $\gamma_s$ spatial **Gaussian random field**.
- **Conjugate prior**: Inverted Wishart for $\Sigma_\epsilon$.
- **Separability assumption**: Residual covariance has Kronecker structure, $\epsilon \sim \text{MVN}(0, I_n(T-1) \otimes \Sigma_\epsilon)$.
- Calibration formula similar to Uni STReg.
Multi STReg assessment

More accurate than Uni STReg when reduced to one dimension when strength can be borrowed.

- Use average of remaining 16 hours of measurements & model outputs on each day to forecast 8-hour (10AM-5PM) peak average.

- Both the RMSPE & RMSFE smaller for the Multi STReg variate model.
Multi STReg assessment

More accurate than Uni STReg when reduced to one dimension when strength can be borrowed.

Multi-STReg makes big changes:

Uncalibrated model outputs versus calibrated (Multi STReg) inferences for selected days.
Multi STReg assessment

More accurate than Uni STReg when reduced to one dimension when strength can be borrowed.

Bayesian Melding does too:

Uncalibrated model outputs versus calibrated (BM) inferences for selected days.
Multi STReg assessment

More accurate than Uni STReg when reduced to one dimension when strength can be borrowed.

<table>
<thead>
<tr>
<th></th>
<th>RMSPE</th>
<th></th>
<th>RMSFE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>multivariate</td>
<td>univariate</td>
<td>multivariate</td>
<td>univariate</td>
</tr>
<tr>
<td>daytime</td>
<td>16.78</td>
<td>18.10</td>
<td>14.44</td>
<td>15.91</td>
</tr>
<tr>
<td>nighttime</td>
<td>12.51</td>
<td>14.19</td>
<td>9.69</td>
<td>9.71</td>
</tr>
</tbody>
</table>

Average of MSEs for spat prediction & temporal forecasting: multi-vs univariate methods.
Conclusions

- Overall Mult STReg best calibrator among the various methods.
- But Uni STReg simpler!
- Correlation between night- & daytime measurements allows strength to be borrowed from the night.
- Approaches show how to infer Policy Related Background levels. (But they vary by hour & region.)

Current work: non-stationary extensions with time dependent coefficients.
References

http://www.stat.ubc.ca/Research/TechReports/tr08.php
http://www.stat.ubc.ca/ jim/

Contact: jim@stat.ubc.ca