A Bayesian Algorithm for Assessing Uncertainty in Radionuclide Source Terms

Dr. Peter Robins
Content

- Inference Algorithm
- NPE 2013 Inference (Simulated data)
- Fukushima Inference (Real data)
- Future
Input data

- CTBTO Radionuclide sensor data
  - 24h integrated Measurements and “Minimum detectable concentrations” for a number of isotopes

- CTBTO Atmospheric Transport Modelling (ATM) dilution fields

- Processed seismic events

- Known nuclear installation locations
Bayes’ theorem

\[ p(\theta| x, M) = \frac{p(\theta|M)p(x|\theta, M)}{\int_{\Theta} p(\theta|M)p(x|\theta, M)} \]

Parameters

Data

Prior

Likelihood

Normalizing constant – can be ignored

Model:

Parameter structure, ATM, decay, sensor response, noise, uncertainty

Continuous values: Probability density – not probability.
Parameterization and Priors – $\theta$, p($\theta|M$)

- **Location**
  - Uniform over earth’s surface

- **Time**
  - Uniform (two weeks)

- **Duration** (log)
  - Uniform (log) with cap on end time

- **Number of nuclides** for each isotope (log)
  - Normal prior (on logN)
    - Huge geometric uncertainty
    - Maximum of $10^{40}$ ($10^{44}$ atoms in the atmosphere)

- **Repeat R times**
  - So far R = 1 or 2 on desktop hardware

- **$m$**
  - Met/ATM model linear interpolator if 2 available
RN Sensor Likelihoods: normal and max value

- Calculate isotope activity $c_{ij}$ from tri-linear interpolation of CTBTO dilution field for each release and decay model.
- Sum activities for multiple releases $c_j = \sum_i c_{ij}$.

$$\ln[p(x_j|c_j)] = -\frac{1}{2} \left( \frac{x_j - c_j}{\sigma_j} \right)^2$$

$$\ln[p(x_j|c_j)] = \begin{cases} 
-\frac{1}{2} \left( \frac{c_j - x_j}{x_j} \right)^2 & c_j > x_j \\
0 & c_j \leq x_j 
\end{cases}$$
Space/Time Information Mixture Model

- Mixture model over 2D space and 1D time
  - 0.5 shared equally between
    - Potential emitter locations
      - 2D normal (isotropic, distance)
      - Extruded uniformly through time
    - Seismic events
      - 3D normal (correlated, strike angle)
      - 0.5 spread uniformly over Earth’s surface and two weeks
- Apply directly to location/time parameters
  - no modelling required
Sampling – densities to probabilities

• Atmospheric transport is an uncertain, diffusive process
  – A “best estimate” of the parameters is inappropriate
  – Uncertainty in parameters must be quantified

• Markov Chain Monte Carlo (MCMC)
  – Output is simple:
    • Count number of sampled parameter sets in regions of interest
      – Ratios (betting odds)
      – Histograms
NPE 2013 Posterior Marginal Location Inference
NPE2013 Activity and Time Marginals

$^{131m}$Xe Log$_{10}$ Bq

$^{133}$Xe Log$_{10}$ Bq

$^{131}$I Log$_{10}$ Bq

$^{133m}$Xe Log$_{10}$ Bq

$^{135}$Xe Log$_{10}$ Bq

Release time / days
Fukushima $^{133}$I Entire Network for 2 Weeks After

Real data (minus JPP38)

More isotope data makes it worse
What has gone wrong?

• Check the modelling provided by CTBTO
  – Adjoint modelling
    • ~2000 collection periods over 2 weeks from ~7 million source locations/times (1°x1°x3hours)
    • Release at collector during collection period, reverse time, reverse wind vectors, keep turbulent diffusion positive, query model on source space/time grid
    • ~2000 model runs rather than ~7 million
ECMWF/Flexpart Integrated Dilution for 1\textsuperscript{st} $^{133}$I detection
NCEP/Flexpart Integrated Dilution for 1$^{st}$ $^{133}$I detection
What is the cause of the modelling error?

- Incorrect modelled wind vector fields as input?
  - Not according to the German NDC
    - Wind fields around the time of the Fukushima reactor accident were well modelled.
    - German NDC forward modelling indicates material does reach RUP60
      - At the appropriate time

- Adjoint models ≠ forward models?
  - But! NPE2013 simulated data was forward modelled
    - Inference using adjoint modelling worked
    - Is the problem at the edge of the plume?
Future work

• Investigate forward/adjoint modelling discrepancy
  – UK Met Office NAME-on-JASMIN service
    • Negotiating a research license

• Can discrepancy be probabilistically modelled?
  – maintain inference speed
    • Cannot afford forward modelling in sampling algorithm
    • Billions of samples (currently lookup and interpolation)
  – provide appropriate uncertainty in modelling to provide uncertainty in reconstructed parameters
Questions?
Basic ATM Uncertainty

• For each sensor with multiple isotope measurements:
  – $s'$ is the true, unobserved dilution

\[
p(x|\theta, M) = \int_0^\infty p(s'|\theta, M(SRS)) \prod_i p\left(x_i\left| s', \text{decay}_i, \theta \right. \right) ds'
\]

• Using a log-normal
  – Geometric uncertainty of 1000
  – Placeholder
    • Investigations underway