Incremental Material Flow Analysis with Bayesian inference

Rick Lupton, Julian Allwood
rcl33@cam.ac.uk

2017 Joint Conference, ISIE and ISSST
25–29 June 2017, Chicago, USA
Outline

1. What is a Bayesian Inference approach to MFA?
2. Example: incrementally mapping global steel flows
3. Discussion
A Bayesian Inference approach to MFA
Peas

$\theta = (X, \eta)$
Uncertain peas

50g of peas please
Uncertain peas

50g of peas please
Uncertain peas

50g of peas please

Weight of peas per pod?
Uncertain peas

50g of peas please

Weight of peas per pod?

Exactly 50g of peas?

50g of peas please

50g of peas please
Inference: updating knowledge

(a) Hypothesis Space

Given a model with one process:

\[ y_0 \rightarrow \text{Process} \rightarrow y_1 \rightarrow y_2 \]

Where mass is conserved: \( y_0 = y_1 + y_2 \)

These are possible hypotheses:

(b) Inference

Observed data \( D \) — e.g. "\( y_0 = 50 \pm 20 \)"

\[ P(\theta|DI) = \frac{P(D|\theta I)P(\theta|I)}{P(D|I)} \]

Prior knowledge \( P(\theta|I) \)

Weak knowledge of \( \eta \)
No knowledge of \( y_0 \)

Updated knowledge \( P(\theta|DI) \)

Weak knowledge of \( \eta \)
Specific knowledge of \( y_0 \)
Sampling from the posterior

Markov Chain Monte Carlo sampling: NUTS

MFA & Bayesian inference

Cencic, O. and R. Frühwirth. 2015.

Summary

To apply Bayesian inference to MFA, we need to:

1. Set up the model structure (possible hypotheses)
2. Relate the observed data $D$ to the model
3. Quantify initial knowledge about model parameters
Example: global steel
Model equations

Process throughputs and flows:

\[ \mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{q} \quad \mathbf{z} = \text{diag}(\mathbf{x}) \mathbf{A} \]

Model parameters and outputs:

\[ \mathbf{\theta} = [q \ \eta \ \phi]^T \quad \mathbf{y} = [q \ \mathbf{z}]^T \]
Observations: link data to model

Flow rates between processes

\[ d_i = z_{JK} + e_i \]

External input flow rates

\[ d_i = q_J + e_i \]

Flows as fractions of input

\[ d_i = z_{JK}/x_K + e_i \]

Gaussian errors: \( e_i \sim \mathcal{N}(0, \sigma_i) \)
Initial knowledge $P(\theta | I)$

"the process yield is between 75% and 85%"

“the flow rate is estimated as 50 tonnes/day with a standard deviation of 10 tonnes/day”

“the process efficiency is between 0% and 100%”

“the flow rate is positive”
Initial knowledge

External inputs $q$: uniform

Process efficiencies $\eta$: logit-transformed normal

Allocations $\phi$: uniform/concentrated Dirichlet
Data sources

Stage 1:

<table>
<thead>
<tr>
<th>Source</th>
<th>Data about...</th>
</tr>
</thead>
<tbody>
<tr>
<td>worldsteel</td>
<td>BF/DR, steelmaking, most products</td>
</tr>
<tr>
<td>Steel Business Briefing</td>
<td>other products</td>
</tr>
</tbody>
</table>

Stage 2:

...
Results: Stage 1
Results: Stage 2
Details of parameters & flows

Pig iron → Oxygen blown furnace

Prior

Stage 1

Stage 2

Ingot casting → Ingots

Cullen et al.

Cullen et al.

Output fraction [%]

Flow [Mt]
Individual samples

Video
Conclusions

- Bayesian inference is effective for incremental MFA
- Data quality assessment is complementary
Conclusions

- Bayesian inference is effective for incremental MFA
- Data quality assessment is complementary

Disadvantages?

- Computational cost (30 mins / stage in this example)
- More information needed? (prior distributions)
Conclusions

• Bayesian inference is effective for incremental MFA
• Data quality assessment is complementary

Disadvantages?

• Computational cost (30 mins / stage in this example)
• More information needed? (prior distributions)

Future work:

• Visualising uncertain results
• Dynamic MFA models
• Model selection
Thank you


Dr Rick Lupton
rcl33@cam.ac.uk
More slides
MFA & uncertainty

MFA & uncertainty

Laner et al (2015): data quality, characterisation, **statistical methods**

- Sensitivity analysis
  - about the model, not tracking uncertainty in the params.
MFA & uncertainty

Laner et al (2015): data quality, characterisation, **statistical methods**

- **Sensitivity analysis**
  - about the model, not tracking uncertainty in the params.

- **Gaussian error propagation**
  - ok if Gaussian approximation is valid
MFA & uncertainty


- Sensitivity analysis
  - about the model, not tracking uncertainty in the params.

- Gaussian error propagation
  - ok if Gaussian approximation is valid

![Diagram showing probability distributions with small and large uncertainty, comparing Gaussian approximation to exact distribution.](image)
MFA & uncertainty


- Sensitivity analysis
  - about the model, not tracking uncertainty in the params.

- Gaussian error propagation
  - ok if Gaussian approximation is valid

- Probability theory
  - use arbitrary distributions, propagate using Monte Carlo simulations
MFA & uncertainty


- Sensitivity analysis
  - about the model, not tracking uncertainty in the params.

- Gaussian error propagation
  - ok if Gaussian approximation is valid

- Probability theory
  - use arbitrary distributions, propagate using Monte Carlo simulations

- Possibility theory
  - use membership functions to represent fuzzy intervals
Likelihood $P(D|\theta I)$

How likely is the model to predict the actually-observed data $D$, if we know the parameters $\theta$?
Likelihood $P(D | \theta I)$

How likely is the model to predict the actually-observed data $D$, if we know the parameters $\theta$?

- Model:
  \[ y = f(\theta) \]

- Observation:
  \[ d_i = g_i(y) + e_i \]
Example: model

Model $y = f(\theta)$:

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X \\ X\eta \\ X(1 - \eta) \end{bmatrix}$$
Example: observation

\[
\begin{align*}
\text{Measuring...} & \quad \text{Likelihood } P(D|\theta I) \\
y_1 & \quad \mathcal{N}(d_1 - X\eta; \sigma_1) \\
y_1 \text{ and } y_0 & \quad \mathcal{N}(d_1 - X\eta; \sigma_1) \cdot \mathcal{N}(d_2 - X; \sigma_2)
\end{align*}
\]
Posterior

\[ P(\theta|I) \]

\[ P(D|\theta) \]

\[ P(\theta|I)P(D|\theta) \]

Uniform prior

\[ \eta \sim \mathcal{N}(0.4, 0.2) \]
Initial knowledge: flow rates

Say we know it **positive** and **less than 100** but nothing else. What do we mean?
Initial knowledge: flow rates

Say we know it **positive** and **less than 100** but nothing else. What do we mean?

- equal probability to equal ranges?
  - e.g. 10–20 kg and 90–100 kg

- equal probability to different orders of magnitude?
  - e.g. 1–10 kg and 10–100 kg
Initial knowledge: allocations

\[ \alpha_i > 0, \quad \sum_i \alpha_i = 1 \]