CIRAIG™

15 years OF SUSTAINABLE COMMITMENT

UNCERTAINTY IN ECODESIGN TOOLS

Pascal Lesage

March 29, 2017

POLYTECHNIQUE MONTRÉAL

Motivation

- Eco-design tools should be quick and easy
- The easiest way to be quick and easy is to use aggregated datasets that one can simply scale and combine
  - **“Aggregated datasets”:**
    - cradle-to-gate inventories
    - cradle-to-gate indicator scores (better!)
- However, uncertainty analysis using aggregated datasets under-estimates uncertainty of results and frankly often makes no physical sense

Outline

- Show why it does not make sense to do uncertainty analysis with aggregated datasets
- Propose two techniques to deal with the method:
  - **An unimplemented method based on the analytical approach**
  - **An implemented method based on presampling**
- Showcase the use of the second method
- Show the extensions of this approach to other parts of the LCA calculation

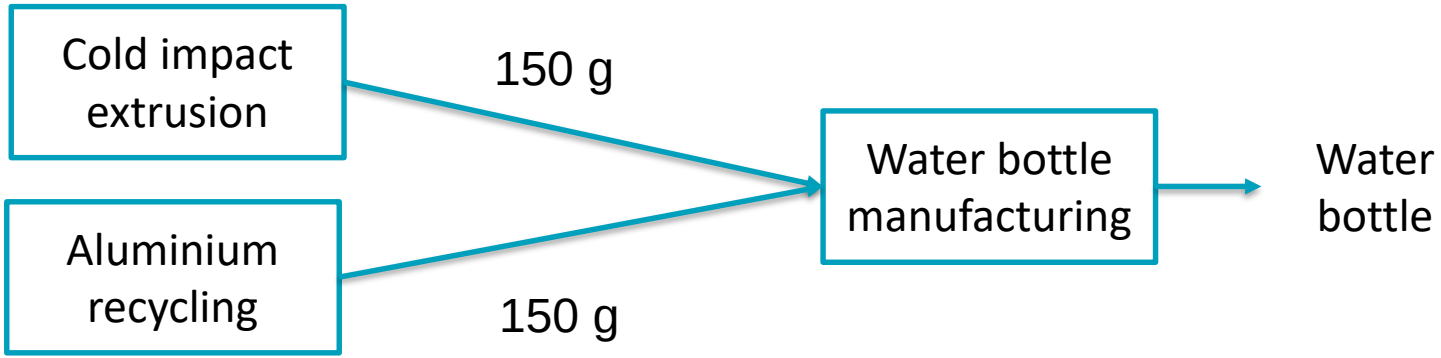
Why aggregated datasets in eco-design tools?

- What's an aggregated dataset?
  - **LCA with disaggregated datasets (gate-to-gate unit processes): need to link unit processes and solve a (large) system of linear equations**

  - **An aggregated dataset is a stored solution to this system of linear equations.**
  - **Stored as an indicator result, it is <u>one number</u>.**

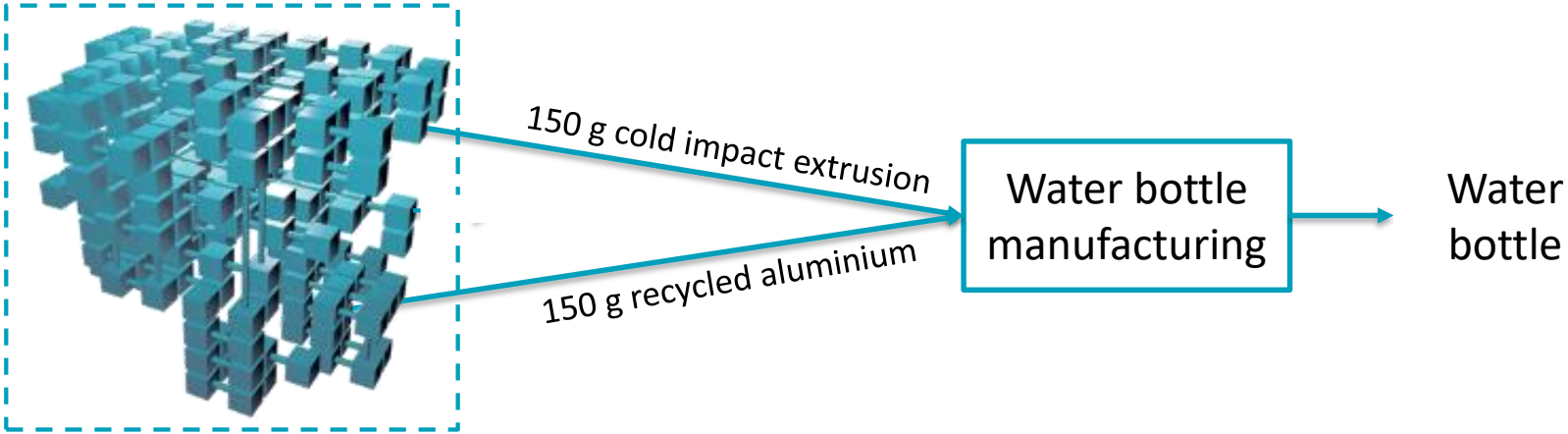
Why aggregated datasets in eco-design tools?

- Take this simple case: producing a water bottle by deformation stroke of recycled aluminium
  - https://www.youtube.com/watch?v=VYCOn-MvGrQ#t=00m51s
- The making of the water bottle is modelled in the simplest possible way:



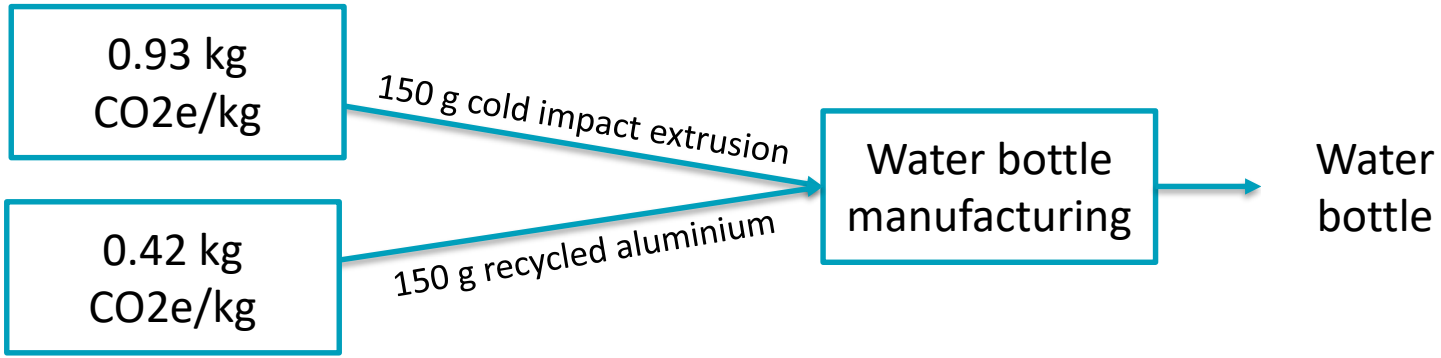
Why aggregated datasets in eco-design tools?

- Take this simple case: producing a water bottle by deformation stroke of recycled aluminium
  - https://www.youtube.com/watch?v=VYCOn-MvGrQ#t=00m51s
- Using unit process data, the model looks like this:



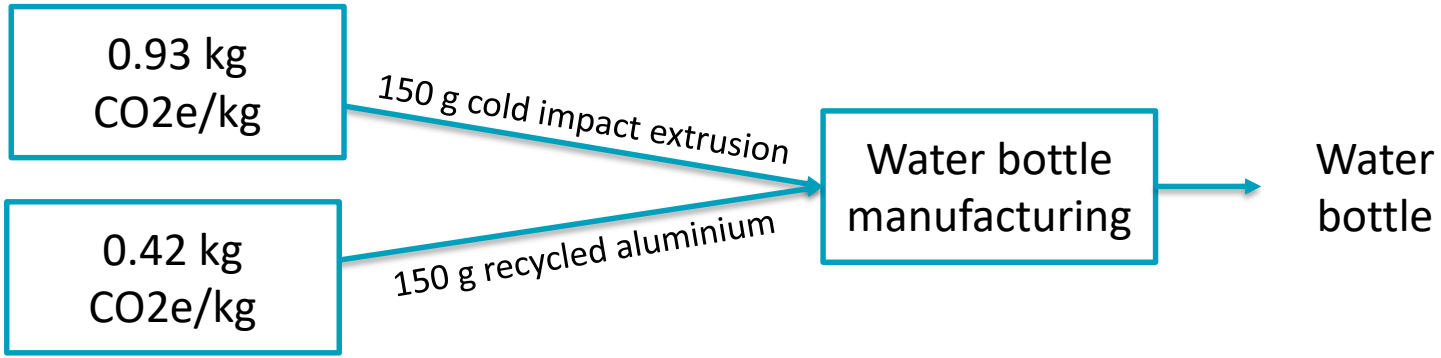
Why aggregated datasets in eco-design tools?

- Take this simple case: producing a water bottle by deformation stroke of recycled aluminium
  - https://www.youtube.com/watch?v=VYCOn-MvGrQ#t=00m51s
- Using aggregated data (say, kg CO2e), the model looks like this:



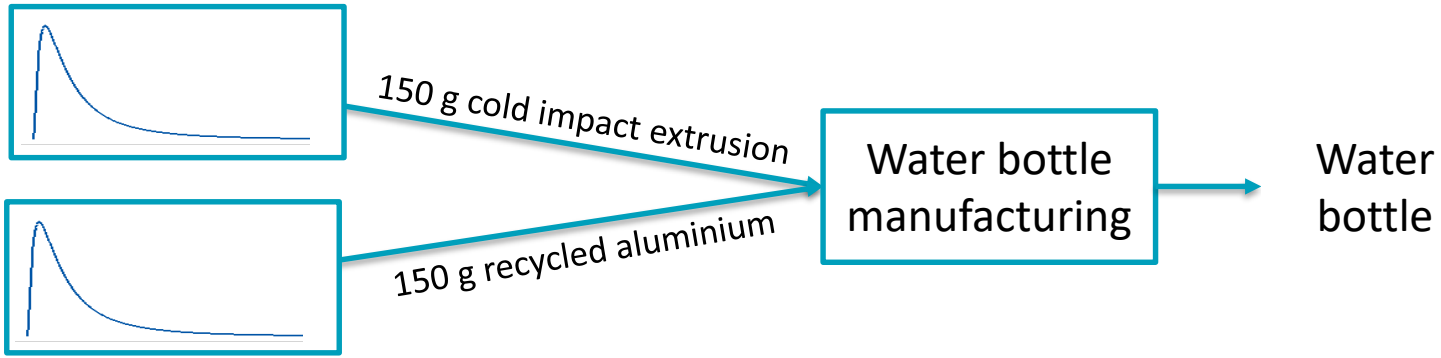
Why aggregated datasets in eco-design tools?

- Knowing that both approaches give the same result and assuming the users of the eco-design tool will *not* want to adapt the unit process data, nor explore it with any detail, the aggregated dataset version is *vastly superior*



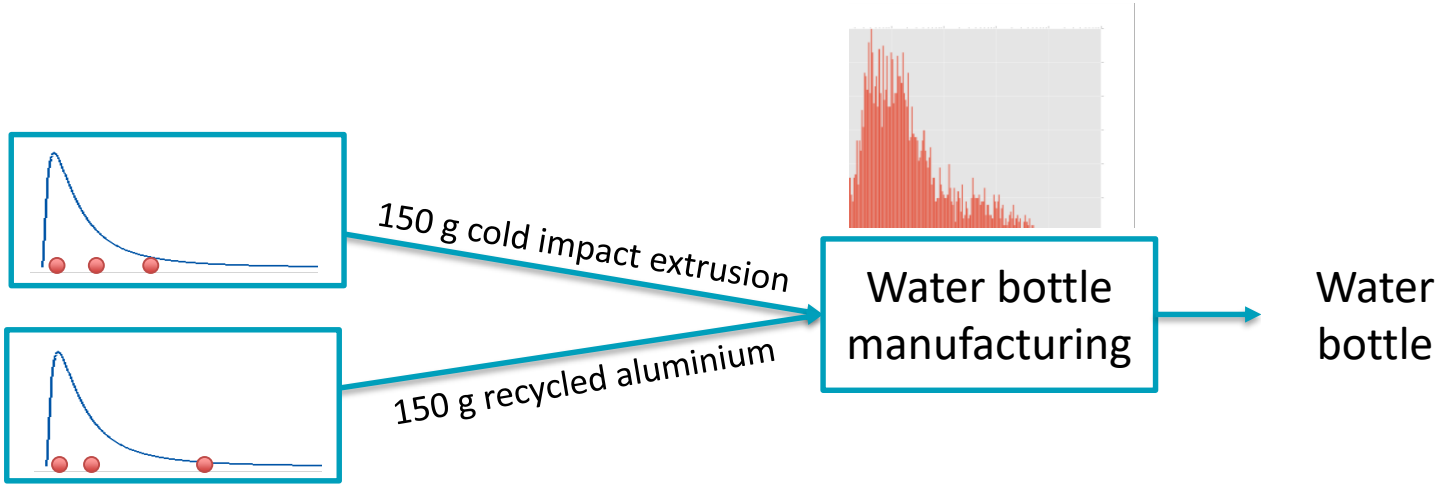
Why not uncertainty analysis using aggregated datasets?

- It is possible to express each indicator result as a probability density function, based on an uncertainty analysis



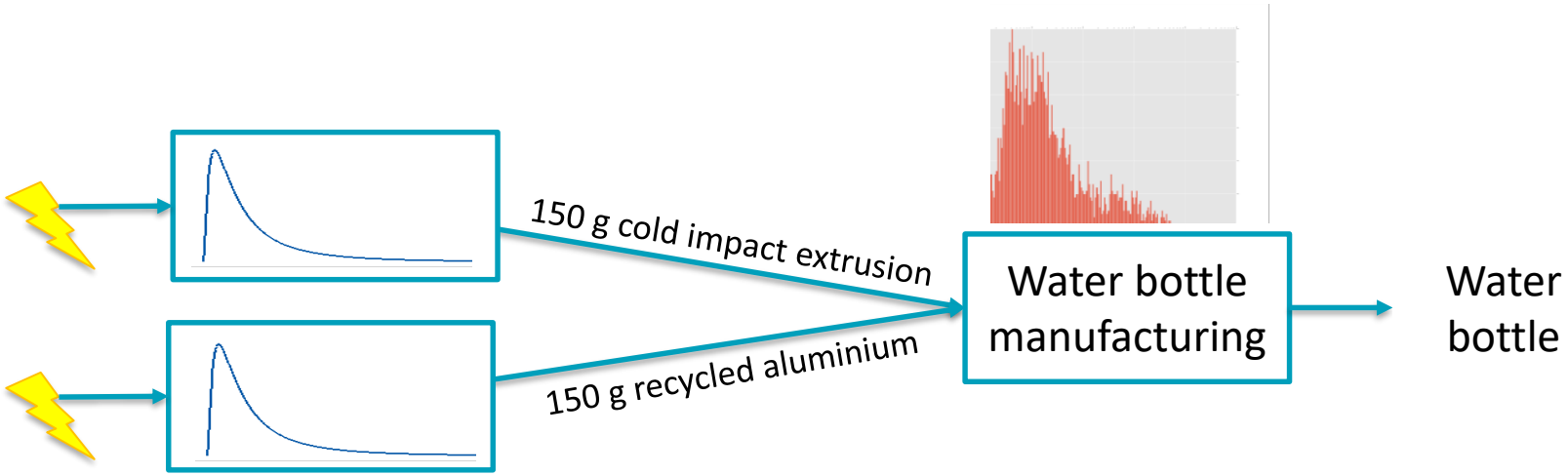
Why not uncertainty analysis using aggregated datasets?

- It is possible to express each indicator result as a probability density function, based on an uncertainty analysis
- And then to sample values from both scores independently to calculate a distribution for the result



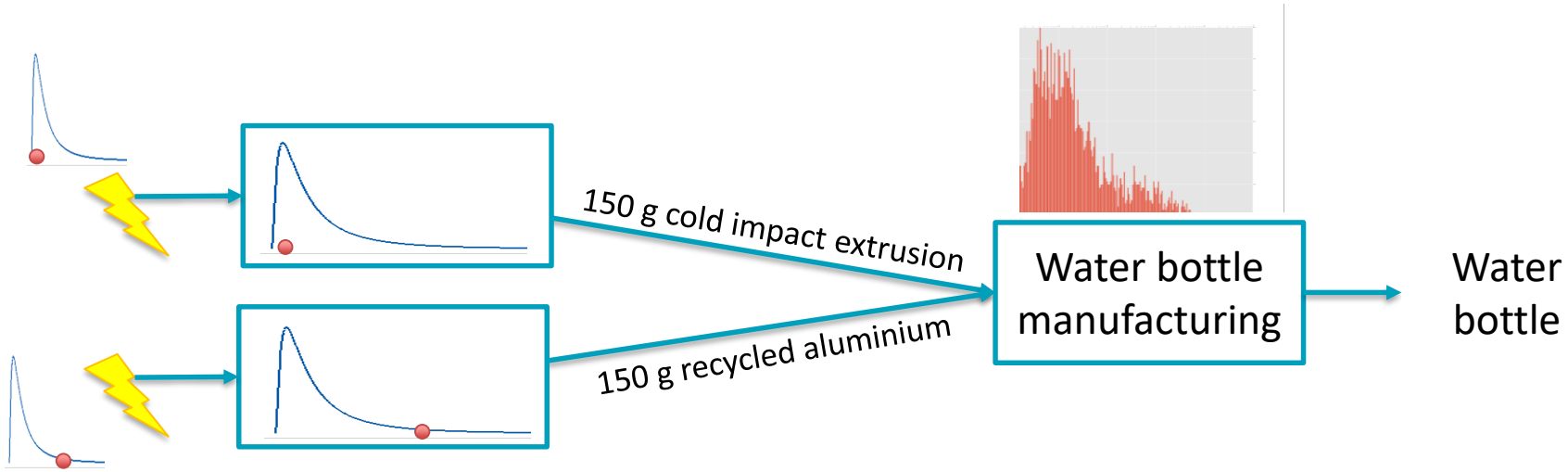
Why not uncertainty analysis using aggregated datasets?

- However, we need to remember that the aggregated data share some common unit processes in their backgrounds. For example, 10% and 50% of the climate change score for recycled aluminium and extrusion come from electricity:



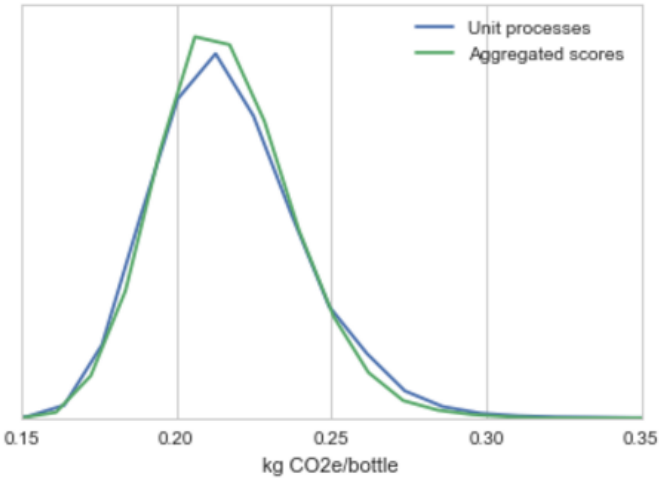
Why not uncertainty analysis using aggregated datasets?

- Values for emissions associated with this electricity is also uncertain, and so the (implicit) indicator result for the underlying electricity can differ widely for a given (independent) sample



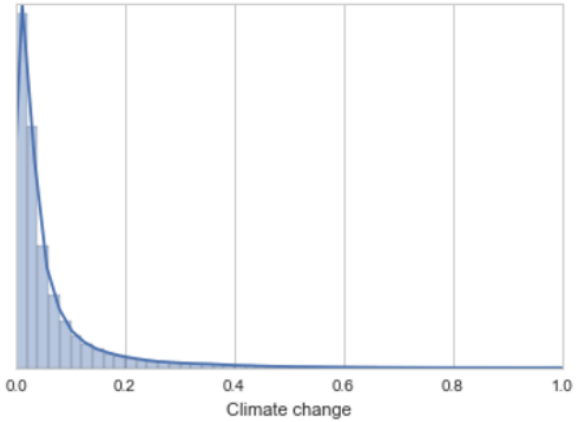
Does that really matter?

- For this made-up example and this impact category: not really.
- The Pearson correlation coefficient of climate change scores between cold impact extrusion and recycled aluminium is 0.4
- The difference in distribution for the result is somewhat negligible
- The “Overlap ratio” (OVL) of the two distributions is 85%



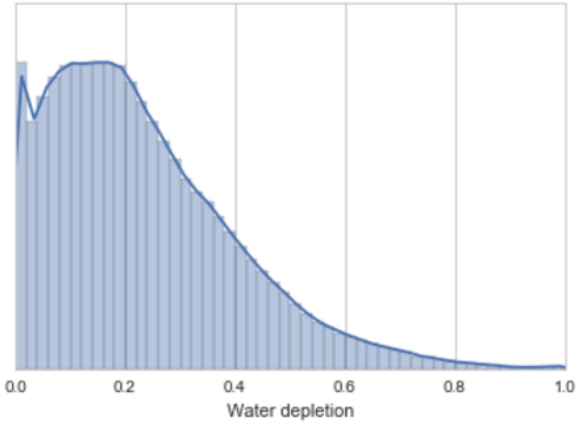
That doesn't seem really that important

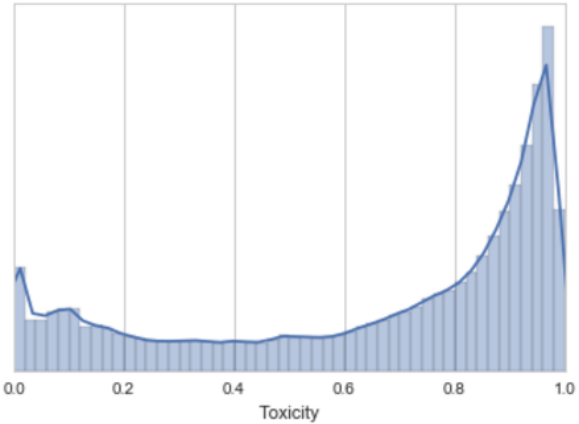
- Well…it depends on the process pair and the impact category
- Here are Pearson correlation coefficients between pairs of ecoinvent 2.2 "product systems" (based on 10000 iterations, over 8e6 pairs)



Not often important

Sometimes important





Often important

OK, so no uncertainty analysis with aggregated datasets?

Wait! There are solutions.

1. Including the (normally excluded) covariance term in the analytical approach

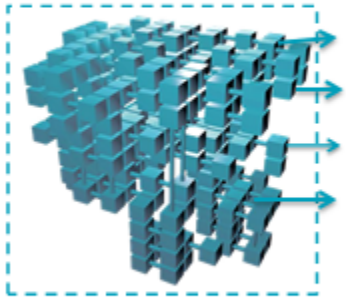
   Not implemented

2. Presampling

   Implemented

Presampling approach – How do we generate the samples

For some indicator *x*



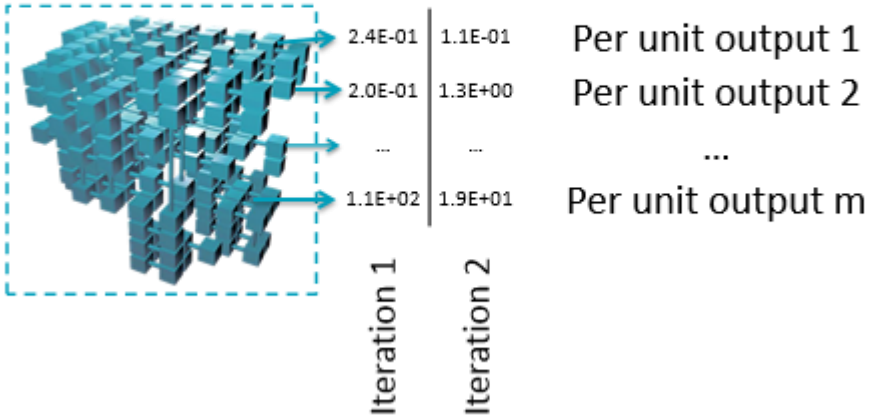
2.4E-01	Per unit output 1
2.0E-01	Per unit output 2
...	...
1.1E+02	Per unit output *m*

Iteration 1

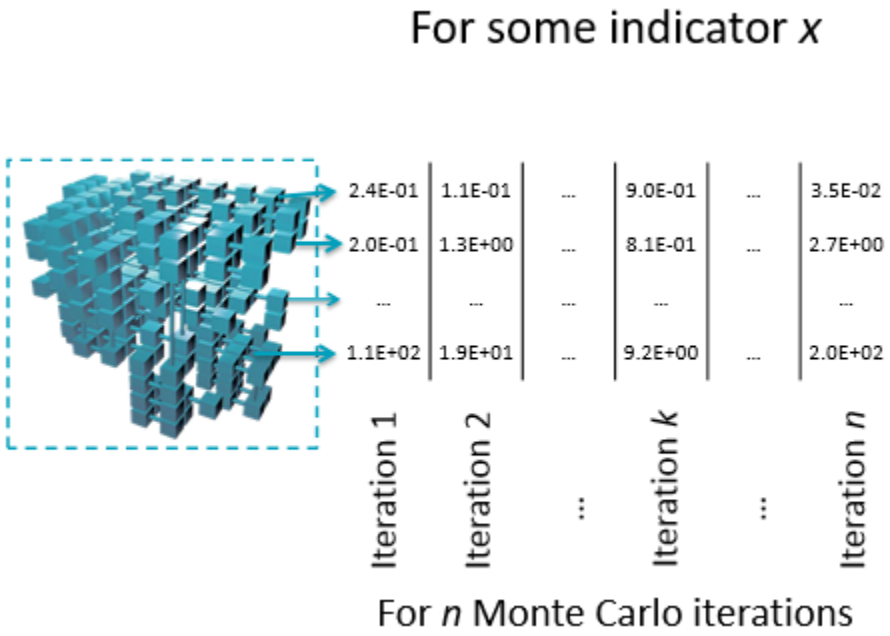
For each unit output from the database

Presampling approach – How do we generate the samples

For some indicator *x*

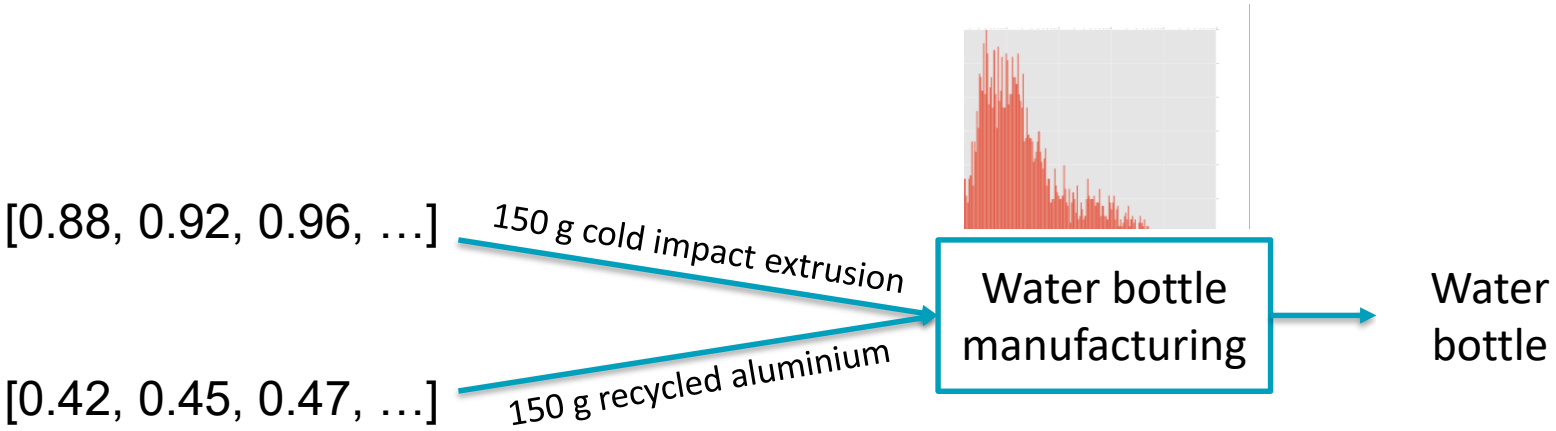


Presampling approach – How do we generate the samples

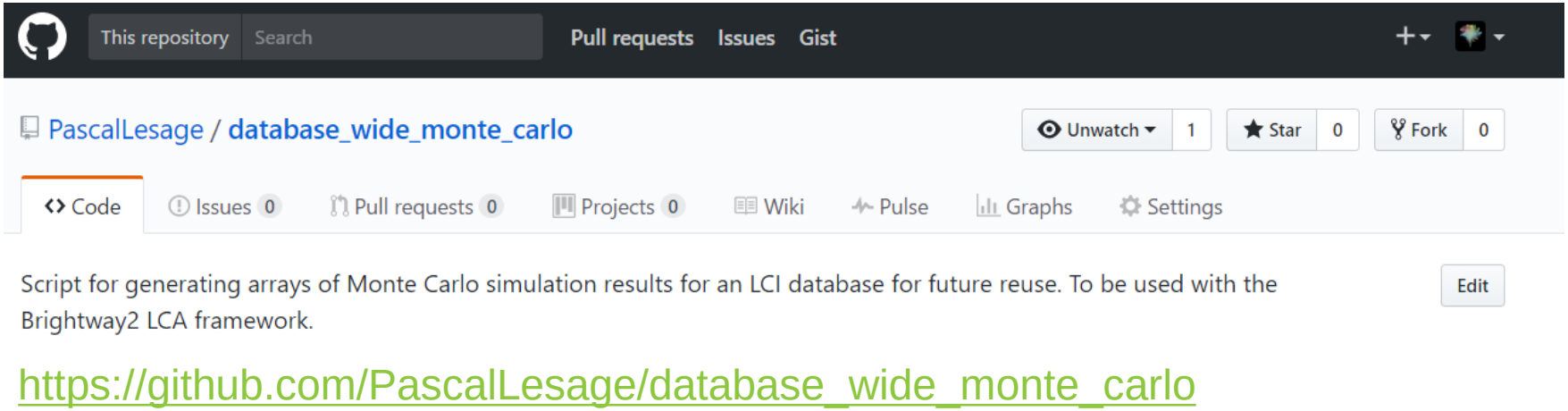


Presampling approach – How do we use the results

- Linear combination of arrays rather than of deterministic results.
- Must conserve the order of the columns!



Presampling approach – Code



https://github.com/PascalLesage/database_wide_monte_carlo

Is this feasible?

- Yes.
  - **Takes a few days to generate/clean samples**
  - **For a 10 000 iteration sample (one product, one indicator), that is about 80 kb**
  - **Only the datasets of interest need to be loaded**

Outlook – this can be used elsewhere!

- The future of LCIA includes uncertainty
- However, characterization factors for a given impact category (or even across impact categories) are not independent!

- Storing arrays of characterization factors rather than characterization factors as PDF would take this into account.

Dedicated partners!

QUESTIONS?