
Brightway-presamples Documentation

Release 0.0.dev

Pascal Lesage

Jun 13, 2019

Contents:

1	Use of presamples in Monte Carlo calculations	3
2	Presamples types	5
2.1	Matrix presamples	5
2.2	Parameter presamples	5
3	Presample directories	7
4	Using Brightway-presamples	9
4.1	Generating presamples	9
4.2	Saving presamples	9
5	Indices and tables	11

This library allows you to use pre-generated values (“presamples”) in LCA calculations. It builds on top of the [Brightway LCA framework](#).

Pre-generated arrays can be used to improve LCA calculations in many ways:

- Seeded pre-sampled Monte Carlo values are reproducible
- Pre-generated static or stochastic values can be generated by complex, non-linear models, capturing system dynamics more accurately
- Pre-sampled Monte Carlo values can capture correlations between parameters, such as between characterization factors, or between input and outputs (e.g. fuel use and CO2 emissions)
- Direct use of population data avoids losses from fitting data to distributions

Brightway-presamples also introduces the idea of *campaigns*, which organize presamples and allow for the systematic exploration of choices.

Use of presamples in Monte Carlo calculations

The Brightway calculation library *bw2calc* already knows how to load and use presamples. Using presamples in a calculation is therefore as simple as passing a list of presamples directories or a *Campaign* object to the LCA class you want to use. Internally, the *bw2calc* library will build the needed matrices for the calculation it is doing, and then replace elements in the matrices with values from the supplied presample arrays.

Warning: Due to the way presamples are implemented and the use of sparse matrices, presamples values must replace numbers *already present* in the matrix to be modified. Presamples should replace existing matrix values, not add new values to existing matrices.

Presamples can be used in both static and stochastic LCA calculations. However, the functions to insert presample values will always pick randomly from the available values. Therefore, if you are using presamples in static calculations you probably want to use presamples with only one possible value for each matrix element.

Presamples types

There are two types of presamples - numbers that can be directly inserted into matrices used in LCA calculations, and named parameter values which can be used in either sensitivity analysis or to generate additional outputs. We call the first type “matrix presamples”, and the second type “parameter presamples”.

2.1 Matrix presamples

Matrix presamples are an array of numbers that can be directly inserted into one of the matrices used in LCA calculations, such as the technosphere matrix. Brightway-presamples allows you to insert values into any matrix, including matrices only used in LCA extensions such as regionalization. When building matrix presamples, therefore, you must not only specify the numbers and where they apply, but also the matrix to be modified. Brightway-presamples contains helper functions to make this easy for common matrix presamples, such as those used in the technosphere, biosphere, and characterization matrices.

2.2 Parameter presamples

Presample directories

Regardless of presample type, all presamples are stored the same way. Each collection of presamples is stored in a directory. More than one presample, and more than one type of presample, can be stored in the same directory. Most commonly, this directory is in the same location as other Brightway2 data, though it doesn't have to be. The directory has the following files:

- `datapackage.json`: A JSON file following the [datapackage standard](#) with the metadata for these presamples. This is the only file that is **required** for a presample directory.
- `{name}.{index}.indices.npy`: A numpy [structured array](#) that gives the matrix row and column index values for the sample data.
- `{name}.{index}.samples.npy`: A numpy array that gives the numeric data to be inserted in the matrices.
- `{name}.{index}.names.json`: A JSON file listing the names of variable parameters.

`name` will be the same for each file in the directory. For each `samples` file (i.e. for each `index` value), there will be either an `indices` or a `names` file.

The integrity of presample directories and their files are checked each time they are loaded.

Using Brightway-presamples

4.1 Generating presamples

TODO

4.2 Saving presamples

Presamples are stored in *Presample directories*, and saved using the function `create_presamples_package`.

4.2.1 `matrix_presamples`

In most cases, `matrix_presamples` is a list of `(samples, indices, matrix label)`.

- `samples` should be a two-dimensional Numpy array.
- `indices` is an iterable with row and (usually) column indices, in the same order and of the same length as `samples`, with the exact format depending on the matrix to be constructed. These indices tell us where exactly to insert the samples into the matrix.
- `matrix label` is a string giving the name of the matrix to be modified in the LCA class. This string is normally one of `('technosphere', 'biosphere', 'cf')`.

Note: If you used the function `TODO`, then your output is already in the correct form.

If the presamples are for the **technosphere** matrix, then `indices` should have the form `[(input id, output id, type)]`. `input id` and `output id` are activity identifiers like `("my database", "my code")`. The `type` is an exchange type, i.e. one of “production”, “technosphere”, “biosphere”, or “substitution”.

If the presamples are for the **biosphere** matrix, then `indices` should have the form `[(flow id, activity id)]`. Both these `id` values should be activity identifiers like `("my database", "my code")`.

Inventory presamples can include values for both the technosphere and biosphere matrices; you can use the `split_inventory_presamples` to split them up.

If the presamples are for the **characterization** matrix, then indices should have the form `[flow id]`. The flow id should be an activity identifier like `("my flow database", "my code")`.

4.2.2 parameter_presamples

`parameter_presamples` is a list of `(samples, names)`.

- `samples` should be a two dimensional Numpy array.
- `names` is a python list of variables names (as strings) in the same order and of the same length as `samples`.

4.2.3 Presamples for custom matrices

TODO: Fill this out and move to a subsection.

Matrices outside the three with built-in functions (technosphere, biosphere, cf) can be accommodated in two different ways. First, you can write a custom formatter function, based on existing formatters like `format_biosphere_presamples`, and add this formatter to the `FORMATTERS` registry:

```
from presamples import FORMATTERS

def my_formatter(indices):
    # do something with the indices
    return some_data

FORMATTERS['my matrix type'] = my_formatter
```

Second, you can also specify the metadata needed to process indices manually. The following arguments can be specified in order, `dtype=None`, `row_formatter=None`, `metadata=None`)

Technical reference

Saving presampled arrays

CHAPTER 5

Indices and tables

- `genindex`
- `modindex`
- `search`