Matching bottom-up and top-down for verification and integration of LCI databases

BY BO P. WEIDEMA, 2.-0 LCA CONSULTANTS, WWW.LCA-NET.COM

International Workshop on LCI-Quality, Karlsruhe, 2003.10.20-21.

Abstract

A combination of traditional bottom-up process-based LCI databases with top-down databases based on industry input-output statistics provides fruitful opportunities for mutual verification. The sum of all processes is verified against the global totals, thus identifying missing data and filling data gaps. The global totals also provide continuously updated normalisation references for the process-based data. At the same time, the process-based data provide much more detail, allowing a breakdown of the industry totals. By integrating the two data sources in one database approach, it becomes possible to avoid data gaps while still providing the necessary detail in process modelling. The integration of bottom-up and top-down is illustrated by two examples: 1) a database linking national agricultural statistics to detailed farm models, providing a comprehensive set of data for Danish agricultural products 2) the use of national material flow analyses and national input-output databases to provide an updated background dataset to which individual process-based data and product life cycles are linked for continuous database improvement.

Keywords

Integration of data of different origin; input-output databases; normalisation

1. Integrating bottom-up and top-down

Traditionally, LCI databases have been produced "from the bottom up", i.e. based on data for specific unit processes, often aggregated or linked into larger processes finally modelling larger parts of product life cycles or entire product LCIs.

An alternative "top-down" source for LCI-data is the national input-output statistics in combination with national environmental accounts and national material flow analyses.

Integrating these two sources of data into the same database provides a number of opportunities, which are further explored in this presentation:

- Combining the advantages of detail and completeness.
- Mutual verification of the two sources of data.
- Continuously updated normalisation references.

2. Combining detail and completeness

2.1 Advantages of the two complementary data sources

Since the bottom-up and the top-down data both model the same reality, they should in principle arrive at the same result.

However, the more specific bottom-up data often have a problem of data gaps, since the detailed and systematic tracing of all inputs and outputs requires a large effort. Besides the more conscious decisions to apply cut-off rules to leave out flows that are considered insignificant, there is always a danger of missing important flows by simple ignorance. From experience, bottom-up LCIs can have data gaps that add up to 50% of the total environmental exchanges [1].

Top-down data, on the other hand, have a problem of lacking resolution, and can therefore not stand alone as a data source for LCIs of specific products. In practice, most top-down data are also indirectly derived from bottom-up data, and therefore rely on the quality of more specific data. Nevertheless, top-down data have a higher degree of completeness, due to the many options for verification of national totals. For example, the national totals for emissions of CO_2, SO_2 and NOX are typically calculated from the detailed statistics of trade in energy carriers, combined with industry-specific emission factors, resulting in highly reliable totals. For heavy metals and several other chemicals, it is also the total trade figures from material flow analyses that allow a verification and reliability of the total emissions. In a few lucky cases, total emissions may also be verified by matching actual levels of pollution, e.g. for N-tot and particle emissions.

When adding up - bottom-up - all known processes within one industrial sector, they should in principle come close to the result arrived at from top-down for the same sector. However, in practice the bottom-up processes often have to be adjusted for omitted or forgotten emissions before the two totals match.

The top-down data thus becomes an important tool for completing the bottom-up data, which on the other hand have the advantage of larger resolution.

Thus, rather than seeing the two sources of data as incompatible, they should be seen as complementary.

2.2 Combining the two data sources in one database

There are few practical problems in combining top-down and bottom-up data in the same database. The traditional LCI-process model as described by ISO 14048 can represent both types of data.

In the context of LCI-databases, the top-down data can best be seen as default processes, which are then broken down into more specific processes by the aid of the bottom-up data. When the bottom-up data represent the entire top-down process, it is possible to make consistent corrections to all the bottom-up processes, which afterwards entirely replaces the top-down process. In the less ideal (but probably more often real) situation, the bottom-up data represent only a part of the top-down process, and the remaining part will have to remain in the database as a residual process.

For example, in modelling the dairy industry, we obtained high quality process specific data for production of the main products milk, yellow cheese, powder milk, butter and spreads. However, this left us with a residual of caseinates, fermented milk, processed cheese, ice cream, whey, lactose and ready-made foods, all dairy products for which we have currently only an average emission factor (calculated from the residual emissions when the emissions from milk, yellow cheese, powder milk, butter and spreads had been subtracted from the original total). As our data collection continues, this residual will be further broken down and will eventually disappear.

3. Examples of verification

3.1 Verifying agricultural emissions

The Danish LCA Food Data Base [2] contains unit process data for 28 representative farm types, together representing the entire Danish agricultural sector. Each farm type is represented by a technical model covering the external and internal (from arable land to stable and from stable to land) turnover of fodder, fertiliser and energy. The technical models are based on standard recommended requirements and technical coefficients, which have a very large empirical basis.

The farm models have been validated at two levels: Internal coherence within each farm type and overall coherence between the sum of farm types and national level input use and production. On the farm level, the validation has primarily been done by checking the coherence between land use, crop yields and livestock production (e.g. the feed needed for the herd matches the home-produced feed plus imported feeds less sold cash crops and the sum of homegrown feeds and sold crops fits the land use).

At a higher hierarchical level the land use has been validated by comparing the sum of area for each crop over all farm types with the national statistics for the same year, e.g. checking that the total wheat area and total wheat yield does not differ more than a few percent from the national statistics. Likewise, the total estimated use of inputs like diesel, fertilizer and concentrated feeds across all farm types have been checked against national statistics. In case of differences that could not be ascribed to an error in a specific type, a general correction factor was multiplied into all types for the relevant input item. This was the case for the nitrogen input, where the sum of the farm models could only account for 95% of the nitrogen purchased.

For energy use, the first run of farm models has only accounted for approximately 50% of the energy purchased according to the national energy statistics. Part of the difference could be explained by combustion of crop residues for heating and a larger fuel use in private cars than initially estimated. However, a significant part of the energy use still remains to be accounted for, which implies that until further information has been obtained, the residual energy purchased has to be allocated over the farm products on a less satisfactory basis (the typical default allocation being the economic value of the output). However, this is still preferable than leaving out this energy use (and its emissions), which would have been the result if the results of the farm models had not been verified against the national totals.

3.2 The Danish input-output based LCA-database

As part of a project for the Danish EPA, we have recently produced an LCA-database covering the entire Danish production and consumption, based on the National Accounting (input-output) Matrices expanded with Environmental Accounts, known as the Danish NAMEA. Imported supplies are modelled on the basis of similar foreign NAMEAs. The basic NAMEA from Statistics Denmark only cover the main air pollutants, based on the annual national CORINAIR reporting. However, based on further national emissions monitoring, material flow analyses and similar national data, typically based on trade statistics and industrial information, it has been possible to expand the coverage to all major emissions as determined by the Danish normalisation reference [3].

Comparing the resulting data to more traditional bottom-up data has so far revealed that transport processes may be more significant than hitherto believed. Since the transports are often spread out over many different products, each with their cut-off, a large part of the total transport ends up being ignored. Similarly, large parts of the emissions from retail trade, repair and maintenance tend to have been left out in raw bottom-up process data.

4. Updating normalisation references

The verification examples in section 3 deal mainly with verification of bottom-up data by top-down data. However, high quality bottom-up data may also in some cases lead to revision of top-down datasets. This may be the case when detailed process data reveal sources of an emission, not hitherto included in the national totals.

In our work with the two databases mentioned above, we found the quality of the detailed models superior in the cases of CH_4 and N_2O emissions from agriculture, which lead us to revise the national totals for these two substances, compared to the national emissions statistics. Similarly, we have found evidence that application of standard emission factors for NM-VOC on available process data result in higher emissions than in the national emissions statistics. While investigating possible gaps in the national statistics, we have for the time being lowered the technical emission coefficients so that the sum of our processes matches the total national value.

These examples illustrate that the integration of detailed bottom-up data into the framework of the NAMEAs allow a continuous updating of national normalisation references for use in LCA. Normalisation references are mainly used when comparing different emissions as part of the impact assessment, but the national normalisation references are fundamentally just an LCI (i.e. a sum of all environmental exchanges) for a national economy.

In 1997, the Danish EPA asked for the Danish LCA-normalisation reference to be updated to the year 1994. This normalisation reference was ready for publication in year 2001 [3] (but due to unfortunate circumstances still not officially published) i.e. giving a delay of 7 years. As a by-product of producing the national LCA-database described above, we updated the national normalisation reference this year (2003) to year 1999 (i.e. a delay of 4 years which is close to the minimum achievable due to the delay in publishing of national statistics). We found that in the 5 years between 1994 and 1999, the normalisation references for some impact categories had been doubled (increases in emissions of nitrogen compounds) for others halved (SO_2-emissions) and for some even reduced with 80% (mainly due to reductions in heavy metals emissions). As long as normalisation references play a significant role in impact assessment, any unnecessary delay in revision of normalisation references should be avoided.

Due to the completeness of the input-output based databases, national normalisation references can be kept as updated as the corresponding national LCA databases.

5. Conclusion

The advantages of combining traditional bottom-up LCI databases with top-down databases based on national input-output statistics have been illustrated. By integrating the two data sources in one database approach, it is possible to ensure completeness while still providing the necessary detail in process modelling.

References

1. Lenzen M, (2001). Uncertainty in IO-based LCI. Pp. 17-27 in Nielsen & Weidema: Input/Output-Analysis - Shortcut to life cycle data? Proceedings of a workshop held in Copenhagen on the 29th of September 2000. København: Danish Environmental Protection Agency. (Environmental Project 581).
2. Nielsen P H, Nielsen A M, Weidema B P, Dalgaard R, Halberg N. (2003). LCA Food Data Base. http://www.lcafood.dk
3. Stranddorf H K, Hoffmann L, Schmidt A. (2001). Normalisation and weighting - update of selected impact categories. København: Danish Environmental Protection Agency. Environmental project. Draft manuscript.