SETAC LCA Workgroup: Data Availability and Data Quality

Rolf Bretz (1998)

CIBA Specialty Chemicals Inc., Consumer Care Division, P.O.Box, CH-4002 Basel / Switzerland (rolf.bretz@cibasc.com)

1 Why another Workgroup?

During the last few years, several initiatives were started to improve the availability and the free exchange of process inventory (LCI) data along the whole life cycle, which is vital for cradle-to-grave LCA work. SPOLD[1] developed a data exchange format [1,2,3] that defines a logical structure for LCI data, and will implement a network, the SPOLD Information system SIS, to provide the technical prerequisites for a facilitated data flow. The Swedish SPINE initiative[2] proposed a standardized data model and a relational database structure that should serve as a "common language" for LCA [4]. These projects were welcomed by industry, consultancy, and academia, and created a strong momentum towards standardization and free data flow; at the presentation of the SIS in November 1997, 12-15 participants volunteered to provide data for the network. Nevertheless, LCI data accessibility is still a serious problem, and data quality / uncertainty estimates are rarely available, with dissatisfactory consequences for the usefulness and credibility of LCA as a tool:

- LCAs of complex systems are still prohibitively time-consuming and costly; time and effort is wasted to duplicate work already done by others on input and auxiliary (background) processes.
- LCI data are presently lacking for wide sectors of industry; this leads to arbitrary cutoff limits, introducing bias, since the less well-documented process always looks favorable.
- Unknown uncertainty margins and questionable data robustness make LCAs useless for decision support, unless the differences between alternatives are so striking that common sense would indicate the same result.

Therefore, the SETAC LCA Steering Committee, in its second three-year cycle of LCA workgroups, decided to address the problem, which is of core interest for LCA applicability. A new workgroup (WG), "Data Availability and Data Quality" was formed at the SETAC Europe Annual Meeting 1998 in Bordeaux; it will investigate methods to:

- improve the **availability** and the free exchange of LCI data
- assess and improve their **quality**
- develop **uncertainty** measures, and establish robustness checks

2 Where are we now?

Data availability and exchange are presently limited by factual and psychological barriers: lack of interest in LCA in certain industry sectors; time and effort necessary to collect LCI data; protection of trade secrets; fear of abuse and misinterpretation of LCA results; concern to attract environmentalists' attention. Analysis of positive examples (such as the initiatives of APME [5] and FEFCO [6]) may give insight in the motivation of data donors and reveal the important driving forces (authorities, customers, eco-labelling schemes). The (real and perceived) dangers of LCI data exchange deserve further study (preferably in case studies) and need to be compared with its benefits. Measures for data protection and trust-building (vertical aggregation over process chains, averaging over several suppliers, involvement of clearing houses / trustees) must be investigated. Finally, the technical means to make existing information easily available (databases, networks, interfaces to LCA software) are not yet mature and need further development.

Data quality has been questionable in many LCA studies of the past. LCIs for complex process trees contain thousands of unit processes with 10^4 to 10^5 or more numerical data elements, derived from a variety of sources with a wide range of precision. Other, more mature disciplines (accounting, engineering, manufacturing/quality control, toxicology) long ago developed error detection and quality assurance techniques (GLP, GMP, accounting standards), which may also be applicable to LCI data.

The relevance and applicability of an LCA is already critically influenced in the goal definition phase [7], by the choice of functional units, system boundaries, and systems for energy supply, transport, waste treatment, etc. Both are also subject to systematic errors introduced by cutoff criteria, omissions, allocation rules (or ways to avoid allocation), the treatment of data gaps, and the use of estimation procedures. Guidance should be given to LCA practitioners on appropriate standard methods to solve the above problems, as well as on the reporting needs (transparency) and precautions to avoid misinterpretation. Such rules would also set the ground for efficient peer review, as stipulated in the ISO 14040 ff. standards [8].

A big problem in the comparability of LCI results arises from the choice of reported interventions: Older studies (such as BUWAL 132 [9]) used to report only relatively few key emissions and coarse measures of raw material consumption (e.g. primary energy), whereas modern LCIs,

especially from the energy sector [10], provide detailed lists of hundreds of parameters (emissions and extractions, sometimes broken down to single elements or molecule species). Contradictory or overlapping sum parameters (e.g. TOC / COD) introduce additional insecurity. Standardization of parameter lists and guidelines for sum parameters are urgently needed to make a meaningful classification/characterization possible.

Data uncertainty has long been investigated, multi-dimensional sets of data quality indicators (DQIs) were developed [11], and stochastic approaches developed [12]. However, the use of DQIs in present-day databases is rudimentary (at best). Error propagation models are missing, and confidence intervals to LCA data are usually not calculated. The practical use of LCAs as decision-making tools depends critically on the existence of reliable uncertainty measures and robustness checks. Practically applicable data quality indicators must be defined, the mathematical models need to be developed into feasible methods (supported by appropriate software tools), and the use of sensitivity analyses should be encouraged by realistic guidelines.

3 What do we want to achieve?

All three aspects, data availability, quality, and uncertainty, were already addressed at the LCANET meeting in late 1996 [13]. Several pertinent projects are ongoing, e.g. the SPINE project, the SPOLD data format and SIS, the enhancements of various commercial LCA software packages to accomodate the SPOLD format, and the data network and DQI implementation for the ESU ecoinventory of ETH Zurich.

The goal of the SETAC LCA workgroup "Data Availability and Data Quality" is to focus these developments, and to find scientifically sound, multidisciplinary, creative solutions to a whole range of problems in LCI data collection. During the planned three-year term of the WG, the workgroup intends to produce one or more guidance documents (to be published by SETAC Europe, as a follow-up to the Code of Practice [7]), which will help practitioners and increase the usefulness and credibility of LCA. In addition, we want to make an input (proposal) to an appropriate (new or updated) ISO standard 1404x (for an overview of the ISO activities, cf. [14]). This will be achieved via two workgroup members who also participarte in the newly constituted ISO WG 6.

The workgroup will draw on a large body of scientific information, and it has already attracted LCA experts from a wide range of countries and affiliations. At our startup meetings in Bordeaux, 24 scientists from many European countries and Japan joined as active members. Besides, there is a growing group of agenda members (so far 25, among them four from the USA), who obtain all pertinent informations and can comment on our work, but who are not obliged to attend all meetings. With this open communication, we strive for a broad consensus in the LCA community on our final proposals.

Our workgroup is very praxis-oriented and comprises representatives of almost all major LCA database owners and software suppliers. Thus it offers a unique opportunity to exchange the technical knowledge accumulated in the various systems, and to benefit from existing expertise. We will merge the experiences gained with SPINE and SPOLD in a joint test, and create a forum for the implementation of our findings in new versions of existing databases and software. Our work will be aligned with ongoing activities of SPOLD: The WG will give an input to the planned SIS database network pilot, and support the subsequent implementation of an international LCI database network.

4 What are the next steps?

In Bordeaux, the WG members selected eight points of major interest, which should give us an effective start and serve as focus for our workplan:

1. Driving forces for data exchange: Authorities / customers / general public
2. Means for trust-building and data protection: Vertical aggregation, data averaging
3. Technical means for data exchange: Databases, networks
4. Technical means for data exchange: Interfaces to existing software
5. Description for scope and goal definition: Energy, transport, waste models
6. Standardized list of interventions: Minimum requirements for parameter list / guidelines for sum parameters
7. Data quality indicators: Practicability
8. Error propagation / confidence intervals / robustness checks: Sensitvity analysis / mathematical models

Small working teams within the WG prepared introductory papers for all eight topics mentioned, which were presented and discussed at our "kick-off" meeting (June 2, 1998, in Arnhem / NL). Some of the themes are closely interrelated, therefore we further bundled the work and created five subgroups, to increase efficiency and manageability of the tasks. Individual workplans for the subgroups were outlined and agreed on by the whole WG.

The in-depth discussion will be done in the subgroups, in small informal meetings or - mostly - by electronic communication. Subgroup interim reports are expected for fall 1998 and will be distributed to all members. The next full workgroup session is scheduled to accompany the SETAC Europe Case Studies Symposium (probably Dec. 1, 1998); thereafter, we will meet 2-4 times per year (next time in Leipzig 1999).

If you are interested to obtain our session protocols with