Evaluating European imports of Asian aquaculture products using statistically supported life cycle assessments

Henriksson, P.J.G.

Author: Henriksson, Patrik John Gustav
Title: Evaluating European imports of Asian aquaculture products using statistically supported life cycle assessments
Issue Date: 2015-11-12

Chapter 7

General discussion and conclusions

7.1 Significant trends and environmental hot-spots in Asian aquaculture production chains

By using the LCA approach presented in the current research, several significant trends could be identified among the environmental impacts caused by different Asian aquaculture production systems. Most trends also persisted using both allocation methods, suggesting that many conclusions could be made with great confidence.

Among the conclusions reached, noteworthy was the importance of feed, with large GHG emissions from capture fishing boats, livestock farming and agriculture. Excess feed was also the major driver for eutrophication, and agricultural pesticides an environmental hot-spot for freshwater ecotoxicity. Together with the significantly lower environmental impacts of the large commercial pangasius farms, this indicates that farm management is strongly linked with the environmental performance of aquaculture production. While this might not be surprising, it highlights the importance of training small-scale farmers, where some of the recommendations in the present thesis should be considered.

In China, the non-integrated tilapia farms in Guangdong had significantly lower environmental impacts than the other systems, with the exception of reservoirs for global warming and ecotoxicity. This trend was again mainly a reflection the eFCRs at the farms, given that feed was the main driver behind most impacts. Promoting and distributing high quality pelleted feeds will therefore be essential, alongside better farm management and feeding practices, to reduce environmental impacts. This also holds true for whiteleg shrimp, where the Vietnamese production systems (eFCR=1.3) resulted in significantly lower emissions than either of the Thai shrimp farming systems (eFCR=1.5). There was, however, no clear correlation between intensity and environmental impacts. Future developments of the aquaculture sector therefore need to consider the consequences of land-use and land-use change, stressing that sustainble intensification is the way forward, but that these practices need to be evaluated, identified and promoted for all types of farmers.

Common environmental hot-spots apart from overfeeding in the different production chains included extensive use of fishmeal (especially from mixed fisheries), dumping of sediments into nearby environments, , landfilling of processing byproducts, high reliance on coal power, the use of certain therapeutants and inefficiently processed byproduct meals. Most of these can, however, be addressed by implementing better policies and farming practices, as well as educating all actors in the aquaculture value chain about the environmental impacts related to aquatic food products.

7.2 Irregularities in current aquaculture LCAs

Several LCA studies of aquaculture systems were already available when the present research commenced (2010). Since, the twelve studies originally reviewed in Henriksson *et al.* (2011) have been accompanied by several additional studies, many that focus on Asian farming systems (Cao *et al.* 2011; Hall *et al.* 2011; Bosma *et al.* 2011; Mungkung *et al.* 2013; Huysveld *et al.* 2013). A commonly observed malpractice among the reviewed studies was the mixing of processes from different background databases, since each database relies upon its unique set of methodological choices. The resulting impacts from different databases are therefore completely incomparable. For example, many studies consulted the processes for the production of fishmeal and other animal derived products in the LCAfood (lcafood.dk) alongside ecoinvent, and/or other LCI databases. The LCAfood database, however, constitutes a consequential LCI database that tries to account for market reactions to changes in demand. Some environmental emissions can therefore come out as negative (e.g. if a product substitutes an environmentally poor product), resulting in emissions completely incompatible to those of the attributional ecoinvent LCI database. This malpractice is partially to blame on software developers that often use the number of available processes as a marketing tool and therefore allow for databases to be mixed without providing inexperienced users with any sort of disclaimer. A simple remedy for this problem is therefore to disable the option of mixing different databases in software, or at least provide warnings to users who do so.

There was also a general lack of transparency into the inventory data used, making critical reviews difficult and reproducing results impossible. This goes against the core of the scientific theory and undermines the academic integrity of most LCA results. As Ioannidis (2012) phrases it: "Efficient and unbiased replication mechanisms are essential for maintaining high levels of scientific credibility". These concerns were amplified by the fact that most studies only present aggregated LCIA results, leaving no insight for reviewers or readers to critically evaluate the decisions made. Poor reporting on primary data also hampers the collective efforts of producing a more extensive LCI data library and obstructs any secondary use of that data (including citations). Since most LCI data have much effort invested into its collection, failing to sufficiently record this data is a waste of resources. More strict requirements by journals and reviewers could therefore transcend case studies beyond their current questionable usefulness (Klöpffer and Curran 2013). This reporting could easily be provided, without compromising the word limit of journals, as supporting information to articles. In the present research a spreadsheet was also developed for this purpose (available at cml.leiden.edu/software/software-quanlci.html), providing an easy way to record and report upon different data references and the dispersions related to them (Henriksson *et al.* 2012c).

Additional impact categories relevant to aquaculture and food production should also be better established within the LCA framework, including impacts on seafloors (Hornborg *et al.* 2012), impacts on food security (Garnett 2014) and impacts on biodiversity (Ford *et al.* 2012). It is, however, important that lifecycle thinking prevails when developing these. Meaning that an impact assessment framework should be applicable to the whole range of different processes causing the environmental damage, including agriculture, livestock, industrial processes, transportation, etc. For those methods that are not relevant to a lifecycle perspective, a risk assessment approach might better be applied as it also takes into account temporal aspects, ecosystems' carrying capacity and synergistic effects. Social life cycle assessments (SLCA) and life cycle costing (LCC) indicators also need to be developed, in order to support more holistic life cycle sustainability assessments (LCSA) (Guinée and Heijungs 2011). The implementation of LCSA might, for example, have provided a more balanced view of small-scale farming in the present research. Throughout the process of expanding the coverage of LCA it is also important to acknowledge that some impacts never will fit into a quantitative framework and therefore need to be communicated alongside LCA results, stressing that decisions should never be based on LCA results alone.

The main methodological topic of debate among the aquaculture LCAs reviewed was the use of different allocation methods. Several studies presented elaborate discussions on the topic (Pelletier and Tyedmers 2007; Fet *et al.* 2009; Avadí and Fréon 2013) and at least two articles have been dedicated solely to allocation in seafood LCAs (Ayer *et al.* 2007; Svanes *et al.* 2011). However, with the level of overall dispersions now quantified it is clear that choices regarding data sourcing often influence results more than the choice of an allocation factor. This becomes even more evident if only relative conclusions are considered (A>B), as significant trends tended to remain coherent across allocation methods. Thus shifting focus towards data quality.

7.3 Data quality improvement options for LCAs

LCA is a tool with inherent demarcation problems, where statistical inference is inadequate and confirmation bias inevitable. Results often build upon large quantities of data and outcomes from complex models supported by insufficient documentation, making the reproducibility of results next to impossible. In the meantime, results are generally presented in a way that induces high confidence, with comparisons of absolute results being commonplace even in scientific literature (Nijdam *et al.* 2012; Tilman and Clark 2014). Strengthening the scientific integrity of LCA studies and adding confidence behind conclusions were therefore identified as areas of priority in the present research.

Starting at the unit process level, we presented a protocol for horizontal averaging of data in **Chapter 3**, where all available datasets could be used and weighted towards a central moment, reducing the influence from data choices and consequently confirmation bias. In addition to this, a method for quantifying overall dispersions defined as the sum of inherent uncertainty, spread and unrepresentativeness was presented. Acknowledging resource constraints as a generic limitation of the data intensive LCA framework, much effort was invested into making the method accessible to the majority of LCA practitioners and understandable to their audiences. In the process of doing so, nomenclature was presented alongside a spread-sheet for calculating overall dispersions.

The importance of defining unit process data is often underestimated, as many unit process parameters act as multipliers during the propagation process, meaning that one erroneous parameter can result in completely skewed conclusions. The generally opportunistic sourcing of unit process parameters is therefore to be blame for much of the discrepancies seen around LCA results today (de Koning *et al.* 2009). This was initially illustrated using the example of soybeans, where we showed that different sourcing of unit process data among studies describing the same system (soybeans from Brazil) resulted in discrepancies among results with up to an order of magnitude (Henriksson *et al.* 2012b). In the meantime, additional layers of complexity (e.g. geographically specific impact categories, effect oriented impact categories, etc.) are constantly being added to the LCA framework (Hornborg 2012; Ford *et al.* 2012), stressing that a general shift from point-values towards distributions is needed.

The moments (central value, variance, etc.) describing distributions, both in unit process data and results, can be expressed in several ways, none of which is "correct". The most common practice in the field of LCA, to my knowledge, is to use the arithmetic mean as the central value. However, when looking across different inventory data sources in more detail, it often becomes evident that mixes of different indicators for the central value are used. This in conjunction with the use of default uncertainties or pedigree estimates fit to a lognormal distribution often results in strange outcomes. For example, assume that two values of 10 are arithmetic means, with one value being assigned a default variance of CV=0.1 and the other value a variance of CV=0.2, both fit to a lognormal distribution. As these values later are propagated into results, the arithmetic means of the two resulting ranges will diverge, as a result of describing a lognormal distribution with an arithmetic mean. If the median instead was used as the indicator for the central value, this deviation would be reduced (but still persist). This as the median is less influenced by extreme values that otherwise can have strong influence on arithmetic means, especially for small sample sizes. The median is also the basis of comparison in non-parametric tests, the only tests that could be correctly consulted in the present research. It is therefore recommended to adopt the median consistently for all LCA parameters and results, and adjust LCA software accordingly. The ultimate strive, however, should be to fit all data to its own distribution and allow for the most appropriate moments to represent this data.

7.4 Features of horizontal averaging and propagation of LCI data

In order to explore how data best could be horizontally averaged and propagated into LCI results, we used the simplified example (relative to the generally complex aquaculture production chains) of Chinese coal power in **Chapter 4**. Initially, the level of horizontal averaging, which historically has been based upon practical classifications such as geographical regions, products produced or production systems, was questioned. It was also shown how these types of classifications often force a diverse set of practices into the same unit process. For example, the existence of flue gas desulphurisation units in coal power plants proved far more influential on acidifying impacts than the capacity or location of the power plant. This demonstrated that spread could be greatly reduced by reclassifying data individually for each dataset, a rationale that also was adopted in the sixth chapter where a unique classification of grow-out farms was defined for each species and country. This feature was even more prominent for other unit processes encountered throughout this thesis work. For example, rice farming in Bangladesh was characterised by two to three different farming seasons (Amon, Aus or Boro). Each of these farming seasons were related to their own sets of farming practices, intensities of irrigation and yields. Consequently the environmental impacts related to the different harvests actually varied more among each other than compared to many neighbouring countries.

Once the unit process dataset had been defined, the LCI results needed to be propagated towards a common functional unit. Several methods for propagating results have been proposed, including Monte Carlo (MC) and first-order Taylor expansion (Huijbregts *et al.* 2001; Imbeault-Tétreault *et al.* 2013; Heijungs and Lenzen 2013). Of these, MC was decided as the most suitable for the pupose of the present research, as it is commonly available in software (Lloyd and Ries 2007) and allow for post-hoc analyses (e.g. goodness-of-fit tests and significance tests) (Heijungs and Lenzen 2013).

7.5 Identifying significant trends using LCA

Given the many methodological limitations and sources of uncertainty identified throughout **Chapter 2** to **4**, the critical question of "which conclusions can be drawn among ranges of LCA results?" remained. By resolving to the concept of dependent sampling, first roughly outlined by Huijbregts (2001) and later explored by Heijungs and Kleijn (2001) and Hong *et al.* (2010), paired results could be generated, allowing for more powerful paired significance tests. However, a prerequisite for applying any significance test is the establishment of a hypothesis, a rare feature in LCA studies. In **Chapter 5** we therefore stress the importance of defining a hypothesis in LCA studies, where significance tests can be used to test the LCA results and reject the null-hypothesis. By only considering the relative differences, one not only reduces the risk of committing a Type II statistical error (failing to assert what is present), but also ensures that identical methodological choices are maintained (with regards to functional unit, system boundaries, allocation, underlying database, impact assessment method, etc.).

The level of correlation of paired results is dependent upon the number of overlapping unit process. Comparing two different pangasius products from Vietnam therefore offers a greater level of correlation, and thus greater resolutions in comparisons, than comparing pangasius fillets from Vietnam with shrimp tails from China. This as a result of more unit processes being shared between the two pangasius value-chains (e.g. feed production, hatchery production, electricity generation, etc.) than between the pangasius value-chain and the Chinese shrimp value-chain.

7.6 Recommendations

7.6.1 Aquaculture

7.6.1.1 Improving feeding practices

Feed was the largest single driver behind most of the impact categories, either through the use of diesel in fishing boats, agricultural pesticides, field emissions or through nutrient effluents resulting from an excessive use of feed and fertilisers. Reducing the amount of feed used should therefore be a priority for the aquaculture sector.

Reducing the inclusion rates of fishmeal in feeds and sourcing fishmeal from sustainable sources are other priorities for lessening the environmental impact of Asian aquaculture chains. This as fishmeal has been associated with many negative consequences (Naylor *et al.* 2009), including overfishing (Pauly *et al.* 2003), physical damage on seafloors (Hornborg *et al.* 2012) and reducing protein availability for the world's poor (Jacquet *et al.* 2009). In the present research we also show that much of the fishmeal sourced regionally is associated with large GHG and eutrophying emissions. Moreover, all shrimp farming systems in the present research, except those in Bangladesh, required larger inputs of wild fish than shrimp produced. This indicates of a net loss in animal protein, pressures on wild fish stocks and competition with food availability. A partial solution for this problem was presented in Cao *et al.* (2015), where we showed that a more extensive use of processing byproducts in fishmeal production could satisfy between half and two-thirds of China's current fishmeal demand (Cao *et al.* 2015).

7.5.1.2 Reusing wastewater and sediments in agriculture

The grow-out site was the hot-spot for most eutrophication impacts as a result of effluents of wastewater and sediments. One of the most efficient ways to deal with these nutrient flows from aquaculture ponds is to reuse them in agricultural fields. This practice may also help to maintain the soil organic carbon on agricultural fields (Boyd *et al.* 2010; Wiloso *et al.* 2014) and reduce the addition of inorganic fertilisers. Treatment ponds and other types of effluent handling are also recommended, but considerations need to be made with regards to gases released from these instalments.

7.5.2 Aquaculture LCAs

7.5.2.1 Choosing a functional unit beyond farm-gate

Most of the aquaculture LCAs reviewed had set their system boundaries at farm-gate with a mass based functional unit of live fish. The consequence of these choices became that byproducts used in feeds (e.g. rice bran or MBM) were allocated large environmental burdens when mass or gross energy content was used as the basis for allocation, while the allocation towards the inevitable fish byproducts that ensue at fish processing remained unaccounted for. Where economic allocation was adopted the situation was the opposite, resulting in products having lower environmental impacts at farm-gate, but not necessarily as processed products (as the value of fillets or tails are much larger than those of the byproducts). Consequently, by choosing a functional unit beyond the processing stage, the discrepancies between the two allocation methods used in this study (mass and economic) were greatly reduced.

7.5.2.2 Land-use and land-use change related to aquaculture

Land-use and land-use change (LULUC) was not explored directly within this thesis. However, the research of Schoon (2013) and Jonell and Henriksson (2014) conducted in parallel to this work stress the importance of considering LULUC when evaluating the lifecycle of aquaculture products. This relates most directly to mangrove deforestation as a result of establishing new aquaculture ponds, but also LULUC impacts resulting from the provision of feed need to be considered. Middelaar *et al.* (2013), for example, concluded that the GHG emissions from land-use change (LUC) resulting from Brazilian soybean farming could be more than six times those resulting from operations.

7.5.3 Life Cycle Assessment

7.5.3.1 Standardising dispersions around LCA results

Producing and processing empirically quantified dispersions around LCA results is today practically doable for all LCA practitioners and should therefore become norm. Many improvements could, however, aid practitioners with this shift. Initially, the LCA community needs to agree upon one consistent nomenclature so that unit process data and results can be communicated in a correct way. Software developers also need to embrace this nomenclature and improve the existing options for including and analysing dispersions. This would include the options for more distributions or statistical moments (skewness, kurtosis, etc.), the propagation of unit process data alongside characterisation factors, paired sampling, multiple allocation factors, accounting for covariance, provide relevant statistical tests (e.g. goodness-of-fit, Wilcoxon test, Friedman test, etc.), R extensions, GPU support and more easily shared models/inventory data. This would also encourage better reporting of data and raise the standard of LCA as a science.

Improving statistical inference with the support of software is also necessary for managing big data in LCA (Cooper *et al.* 2013). The adoption of big data would reduce the incidence of flawed parameters in LCIs and ultimately harmonise results. It could also come to support long-term datasets and help to update parameters in real-time (Xu *et al.* 2015). The protocol developed presented in **Chapter 3** could be used for the integration of big data into LCA, as working with weighted mean based upon a pedigree approach could assure more objective representation of different parameters (Xu *et al.* 2015).

7.5.3.2 Structuring LCI models to address hypotheses

When adopting dependent sampling, the structuring of the unit process dataset becomes increasingly important. Given that only distributions in unit processes shared by production chains can be dependently sampled, the LCI modelling structure will influence the level of correlation among results. For example, **Fig. 7.1** demonstrates a hypothetical scenario where burning of diesel in two different fishing fleets has been separated into two country specific unit processes. Consequently, the distributions in the two unit processes can only be independently sampled. If the unit process dataset instead was modelled according to **Fig. 7.2**, the national label on emissions might be lost, but dependent sampling prevails. Similar thinking could be applied to all unit processes that rely upon fairly generic data, which also often is related to large spread (e.g. combustion of diesel in unknown engine, wastewater from processing plants, transportation distances, etc.). Constructing unit process datasets and LCI databases accordingly would therefore reduce relative uncertainties, even if absolute uncertainties might increase (by the use of more generic unit processes).

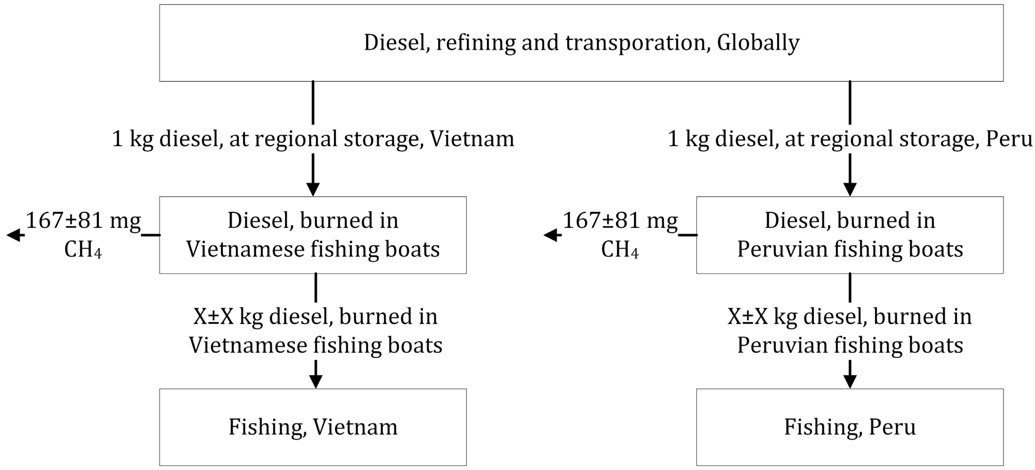

Fig. 7.1: Example of a unit process dataset using separate unit processes for the combustion of diesel in fishing boats.

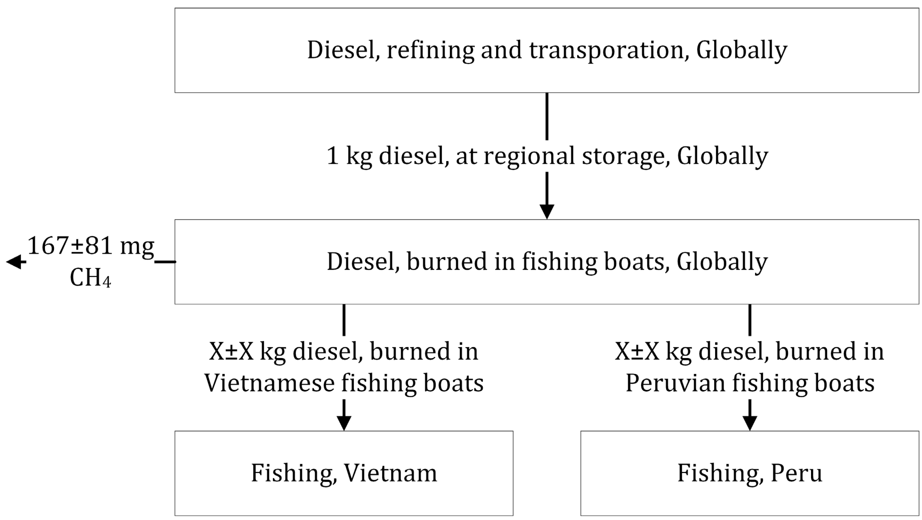

Fig. 7.2: Example of a unit process dataset using only one unit process for the combustion of diesel in fishing boats.

7.5.3.3 Achieving mass balanced LCI models

Since the appearance of computers, mathematical modelling has become the answer for evaluating most of our environmental concerns. Over time, these models have become increasingly complex, leaving ever less room for critical evaluations of the predicted outcomes (Pilkey and Pilkey-Jarvis 2007). LCA is a prime example of such an environmental modelling tool where one flawed parameter or erroneous decimal point can skew conclusions. Striving towards mass balanced LCI models could therefore greatly reduce the risk of such mistakes and logically makes great sense (inputs=outputs). Resolving the many challenges related to this (e.g. chemical reactions within processes) and providing software to support mass balanced models is therefore encouraged.

7.5.3.4 LCA as a science – confirmatory or exploratory

Exploratory research sets out to identifying indicators, rather than being a pathfinder (Tukey 1980). Confirmatory research, on the other hand, aims at identifying significant trends in stochastic environments. While the prior may provide highly valuable information, it does not do it with the same conviction as the latter. Throughout the present thesis, much doubt was shed on the confirmatory use of LCA results, but a more scientifically rigid approach to LCA was also presented. By adopting the suggested approach, LCA practitioners are allowed to achieve statistically supported conclusions, with a reduced chance of committing Type II statistical errors. It is, however, my personal belief that LCA should be used for both purposes, depending upon the goal of the study; where hot-spot analyses and system mapping may help formulate hypotheses for follow-up confirmatory LCAs. I also believe that methodological alternatives add confidence to LCA results, rather than erasing comparability. As was shown, absolute results are irrelevant, so fewer resources should be invested in seeking consensus on methodological choices through operational guidelines, PCR standards, etc. LCA results will always remain incomparable across studies and between LCA practitioners. Exploratory LCA case studies should therefore avoid comparisons with other studies apart from maybe building consensus around environmental hotspots. Finally, it is important to highlight that LCA is not a tool created to save individual species or unique locations, it is a tool crafted to steer societies (not individuals) towards more sustainable choices and actions.