

## Impactmap: Designing Sustainable Supply Chains by Incorporating Data Uncertainty

**Mark Fuge** University of California, mark.fuge@berkeley.edu

**Katherine McKinstry** University of California, mckinstry@berkeley.edu

**Kevin Ninomiya** University of California, kevninomiya@berkeley.edu

**Abstract.** Incorporating uncertainty is critical to developing robust decision support tools for sustainable design. Currently, designers and consumers have few tools that compare the ranges of environmental and social impacts among potential supply chain options. Methodologies such as Life Cycle Assessment (LCA) often provide point estimates of the impacts of materials and manufacturing operations on unit bases, but they are not as well suited to express uncertainty associated with many supply chain choices. This work presents a decision support tool that assesses the sustainability of supply chains while incorporating uncertainty. The core contribution is a methodology for assimilating uncertain impact data such that designers or consumers can compare the impacts of a product's possible supply chains. This is possible even if data are missing, through the use of a form of Uncertainty Propagation that factors in data from similar supply chains. A web-based decision support tool combines impact information from disparate sources and visualizes the resulting uncertainty in an intuitive way. We discuss several application areas, as well as techniques for determining when our methodology might be limited.

**Introduction.** Where and how a product is made has significant impacts on its triple bottom line, or economic, environmental, and social impacts. Much of the world's manufacturing is moving from developed to developing countries where rules on occupational health and environmental emissions are less stringent. Upfront design decisions determine the complexity of assembly and material selection, affecting where and how something will be made. Despite this importance, there are a limited number of tools that can help designers make informed decisions regarding their product supply chain amidst uncertain information. A case study of a laptop computer recycling supply chain demonstrates how the methodology handles and presents real, imperfect data to help the user make optimal supply chain choices. This methodology enables several applications: 1) improving LCA estimates by better utilizing crowd-sourced data, 2) reducing LCA estimate uncertainty with targeted data gathering, and 3) robustly computing low-impact supply chain choices using uncertain LCA data. Overall, this combined approach merges LCAs from various sources to create initial estimates, despite potential missing data—a task none of the comprising estimates could do by itself.

**Related Work.** Related research falls under three areas: 1) LCA tools that incorporate tradeoffs amount Triple Bottom Line factors, 2) the use of quantitative Data Quality Indicators (DQI) for

*Proceedings of the International Symposium on Sustainable Systems and Technologies* (ISSN 2329-9169) is published annually by the Sustainable Conoscente Network. Melissa Bilec and Jun-ki Choi, co-editors. ISSSTNetwork@gmail.com.

Copyright © 2013 by Mark Fuge, Katherine McKinstry, Kevin Ninomiya. Licensed under CC-BY 3.0.

### Cite As:

Impactmap: Designing Sustainable Supply Chains by Incorporating Data Uncertainty. *Proc. ISSST*, Mark Fuge, Katherine McKinstry, Kevin Ninomiya. <http://dx.doi.org/10.6084/m9.figshare.810431>. v1 (2013)

propagating uncertainty, and 3) visualization techniques for encoding complex information about supply chains. This work synthesizes elements from each of these areas into a decision support tool, highlighting some promising research directions at their intersection.

First, while LCA research has largely focused on modeling environmental burdens, sustainability research increasingly recognizes that factors such as social and economic sustainability are needed to make sustainable decisions. In the context of supply chains, Hutchins and Sutherland (2008), Yakovleva et. al. (2011), and Nagurney and Nagurney (2010), all provide frameworks for aspects of the triple bottom line. Our work extends this prior work in Triple Bottom Line (TBL) indicators by providing a method for visualizing those tradeoffs and the associated uncertainty.

Second, a central problem for supply chain LCA is that data sources are diverse and of varying quality. Research in quantitative Data Quality Indicators (DQI) attempts to derive quantitative estimates of uncertainty using that data quality (Lloyd and Reis 2007). All of these approaches derive mathematical forms that allow them to translate subjective DQI indicators into numerical estimates; e.g., hybrid Monte Carlo simulations (Wang and Shen 2013) or Bayesian Hierarchical Models (Neuman et. al. 2012). Though both utilize DQIs, unlike traditional LCAs which are concerned with accurate measurement of existing products and processes, ImpactMap is aimed at earlier stages when a designer is selecting potential supply chains or processes. In these cases, the focus is not on the accuracy of the resulting LCA, but rather on identifying potential best or worst case scenarios to direct further research efforts and gather more information.

Lastly, a number of online tools provide visualizations related to sustainability and supply chains, but few combine this information with TBL and uncertainty data. The most common visualization of supply chains is through a 2D map projection; i.e., Sourcemap (2010) and Patagonia's *The Footprint Chronicles* (2012). For more specific TBL indicators, several sites offer dashboards that aggregate and display impact information. For example, The Social Hotspots Database (2010) aggregates social impacts, and Earthster (2012) aggregates environmental impacts. Sustainable Minds (2012) and EIO-LCA (2008) both aggregate TBL indicators across industry sectors, but they do not explicitly factor in supply chain decisions or uncertainty across that chain. ImpactMap combines the TBL indicators, uncertainty modeling, and visualization strategies of prior work which provides a more informative view than the sum of its parts.

**Methodology.** Impact Map allows a designer to compare different supply chain choices based on their own user-selected triple bottom line criteria. This tool accommodates data from a range of sources and displays the variance based on this quality.

**Data Collection and Categorization.** Since TBL impacts data are typically found in different databases, a standard data collection protocol with four categories was established:

- Question Selection- What type of information was being sought?
- Unit Selection – In what units is the information being stored? (e.g. USD/hour, kg CO2 eq.)
- Geographic Specification – Where is this data relevant?
- Level Selection – How process-specific is this data?

Quality of data was rated based on three criteria: (1) Acquisition method (was the data measured, or an estimate), (2) Independence of data supplier (was data provided by independent source or interested party) (3) Temporal correlation (is this data taken within the last few years or is it old) (Junnila & Horvath, 2003). This data quality was later used to indicate the level of certainty, or variance, associated with the overall sustainability results obtained. Lastly, a precedence hierarchy was established to show the precision of the data by location and process (Fig. 1), allowing us to quantify how closely a data source matches the location and process we are interested in. For example, when classifying worker wages, finding an average

wage for the entire recycling sector is more general than finding wages for a worker working in an e-waste recycling plant performing the task of collection and sorting.



Figure 1: Data Approximation Levels. If data is not available for a given facility or process, we approximate it with less accurate data from outer levels, such as the general industry or country.

**Uncertainty Propagation.** Given a particular set of data, our goal is to determine how possible supply chains compare to each other, and what changes in the chain are likely to have a large impact. To do this, we normalize each of the triple bottom line impacts on a 0-100 scale and use data quality to estimate the uncertainty in that score, where 0 is the worst possible supply chain choice in the database, and 100 is the best. The TBL impact scores are composed of a weighted sum of sub-indicators that have been suitably normalized; for example, social impacts might be composed of estimates of child labor rates and living wage conditions. Each of these sub-indicators has its own associated uncertainty, and propagates that to the overall impact scores. The approach to handling this propagation centers around three main assumptions:

1. The sub-indicators for a particular location or process are modeled as normally distributed random variables whose variance is controlled by the quality of data.
2. If indicator data specific to a particular process or facility are missing, the model substitutes more general data in its place, at the cost of additional variance.
3. If certain sub-indicators; e.g., child labor rates, eco-toxicity levels, etc. are more or less important to users, they can use a weighted linear combination of each indicator.

**Modeling Uncertainty in Individual Indicators.** The first element of uncertainty in the impact of a supply chain lies in estimating a particular sub-indicator value for a process and location; e.g., CO<sub>2</sub> Emissions for a silicon etching facility in Fremont, CA, USA. This paper treats each of those data points ( $d_i$ ;  $i \in \{1, \dots, D\}$ ) as a normally distributed random variable where  $\mu(d_i)$  is the value of a sub-indicator at  $d_i$  normalized to  $[0, 1]$  by the max and min values over all data points for that indicator. Constants  $max_{DQ}$  and  $min_{DQ}$  are the max and min data quality numbers.

$$\mu_{ind}(d_i) = \frac{d_i - \min_{d_j} \mu_{ind}(d_j)}{\max_{d_j} \mu_{ind}(d_j) - \min_{d_j} \mu_{ind}(d_j)} ; \quad v(d_i) = \alpha \left( \frac{\max_{DQ} - DQ(d_i)}{\max_{DQ} - \min_{DQ}} \right) + \varepsilon \quad (1)$$

The first key assumption in this work lies in the definition of  $v(d_i)$ , which is a linear function of a point's data quality  $DQ(d_i)$ . The  $\varepsilon$  parameter defines the minimum noise attainable, and the  $\alpha$  parameter defines a penalization term for poor data quality. These constant can be set to mirror the DQI bounds given by prior work, or might be estimated from data.

**Borrowing Data from Related Processes.** If sub-indicator data is missing for a supply chain process and location, we would adapt information from related locations and processes to approximate the estimate. Consider a process in a facility as being the deepest in a series of levels, shown in Fig. 1. If no information is available for that facility, but the average for that process across the city in which the facility is located is known, this gives a noisy estimate for that facility. If city information is unavailable, we propagate up to region and country level estimates. An analogous procedure happens for process, industry, and economic sector data. As the available indicator data becomes less related to the specific process and facility, we increase the variance using a *variance multiplier*  $\beta$ , which has the following linear form:

$$\beta = 1 + \gamma \cdot \frac{loclevel + processlevel}{\#loclevels + \#processlevels} ; \quad \tilde{v}_{ind} = \beta \cdot \bar{v}_{ind} \quad (2)$$

where *loclevel* and *processlevel* refer to how many levels away from the bottom the data is from (e.g., facility = 0, city = 1, region=2, etc.), and *#loclevels* and *#processlevels* refers to the maximum number of levels (3 for location, 2 for process). The constant  $\gamma$  sets the strength of penalization for the approximation, and can either be set manually, or estimated from data.

**Combining Indicator Estimates.** At this point, all sub-indicators for the entire supply chain have been estimated individually as normal random variables. For the entire supply chain, we use a weighted linear combination of these estimators, where the user can input values for how important various sub-indicators are. These weights are then normalized ( $\sum_{i \in sub-ind} |w_i| = 1$ ), and the summary indicator mean and variance is given by:

$$\mu_{ind} = \sum_{i \in sub-ind} w_i \cdot \bar{\mu}_i \quad ; \quad v_{ind} = \sum_{i \in sub-ind} w_i^2 \cdot \tilde{v}_i \quad (3)$$

This assumes that the error in each part is uncorrelated to the other parts.

**Map Visualization and Interaction.** ImpactMap encodes supply chains through a 2D map that visualizes both the structure and impacts of the supply chain, as well as data about the supply chain impacts. As shown in Fig. 3(a) the map can represent data at different levels of fidelity: country, region, city, and facility. The colors for the map are derived from the weighted indicator factors calculated in Eqn. 3, which are linearly transformed into a red (worse) to green (better) color map. Clicking on a location creates a comparison view, shown in Fig. 2(b), where the color scale shifts to display the difference in means normalized by the selected location's uncertainty:

$$S(l, l') = \frac{l - l'}{\sigma(l')} \quad (4)$$

where  $l$  is any of the possible alternative locations,  $l'$  is the selected reference location, and  $\sigma(l')$  is the standard deviation of the impact value for the reference location. This scaling makes it easier to determine net positive (green) vs. net negative (red) changes, and it incorporates the data uncertainty to help identify choices that will make a meaningful difference. For each impact, a gauge with 0-100 scale is provided where 0 represents the worst possible choice of supply chain and 100 the best. The grey region represents a 95% confidence interval.

**Indicator Components and Weightings.** If a particular user desires more information, he or she can click on the name of a particular indicator and a set of additional information will drop down beneath that indicator, as shown in Fig. 3(c). In the cases where a user has different preferences regarding which indicators they think are important, they can adjust the weighting of each indicator as shown in Fig. 4(b). Negative values represent indicators where a higher amount of the indicator (such as emissions) is undesirable.

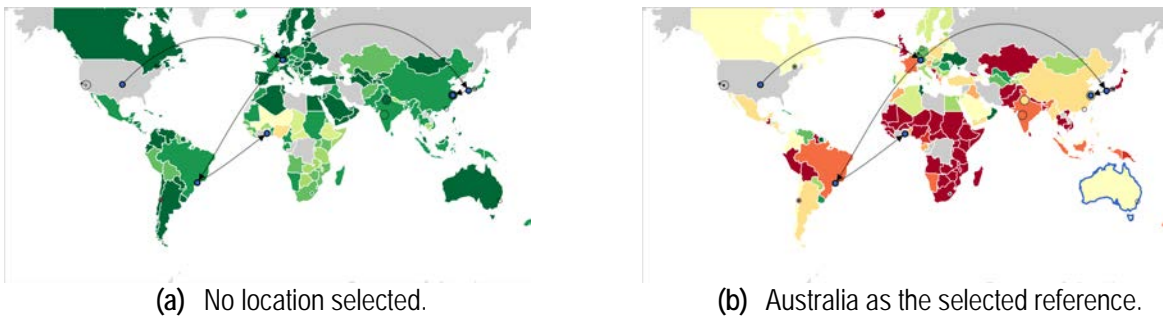


Figure 2: Comparing Reference Locations. Clicking on a particular country renormalizes impact estimates with respect to that location.

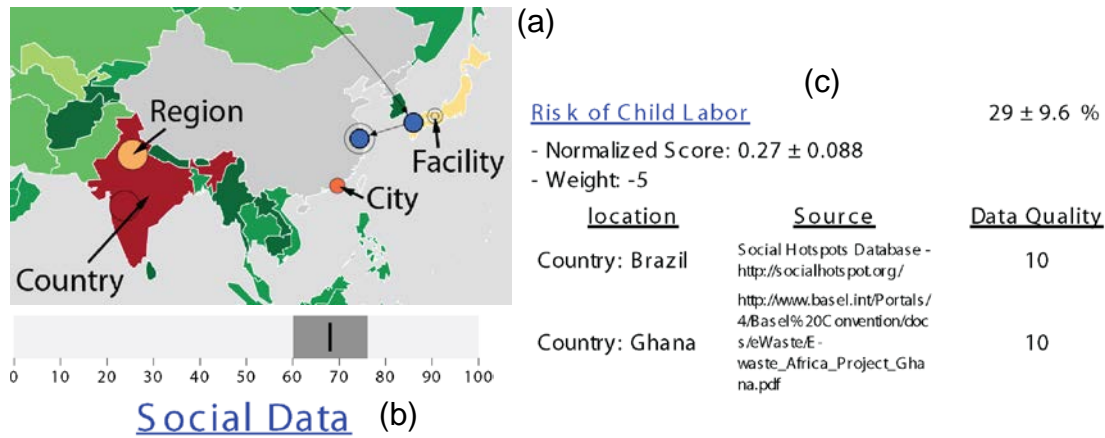


Figure 3: Elements of ImpactMap. (a) Indicator information at different granularity. (b) Summary gauges for each indicator. (c) Detailed source information for further investigation.

**Case Study.** An example case study was conducted by selecting a laptop recycling supply chain, and used ImpactMap to assess possible supply chain options. An example of a recycling chain is shown in Fig. 2(a), which has fairly good social impact performance, but relatively poor environmental performance. Of the supply chain locations, one can see that the processing location in Japan has the lowest environmental score across countries of interest (Fig. 2(a)).

One can observe relative changes in environmental indicators by selecting alternative processing locations within Japan such as Kato, Japan. It has significantly better environmental performance, along good general performance in South Korea, though no specific facilities for that process exist in the database. This provides avenues for future identification of suppliers. Figure 4 demonstrates how these conclusions might change when shifting the weight of the environmental focus. If solely considering Greenhouse Gas (GHG) emissions by giving zero weight to other sub-indicators (e.g., water toxicity), then Japan far outperforms surrounding options (Fig. 4(c)). Depending on the priorities of the company, ImpactMap can be used to explore different combinations of economic, social, and environmental indicators.

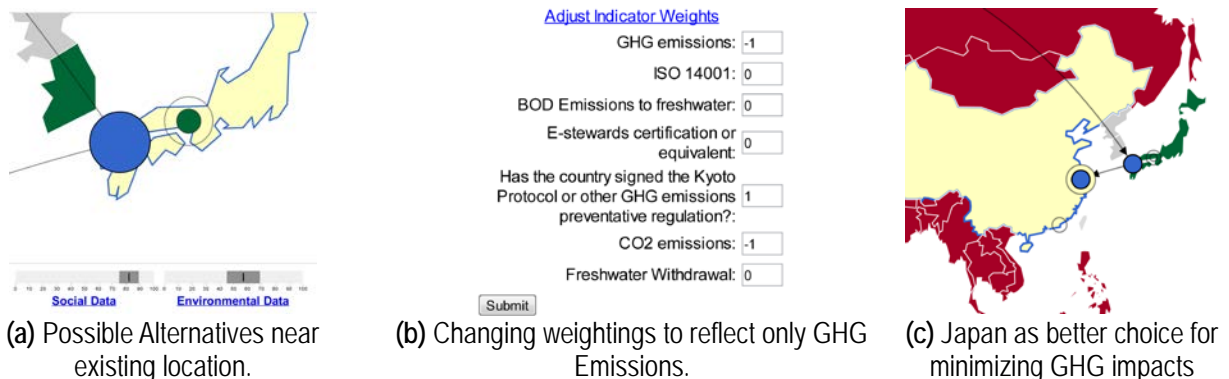


Figure 4: Selecting Alternate locations. Under uniform environmental weightings, we may choose a different processing facility in Japan or South Korea to improve environmental impacts.

**Discussion and Conclusions.** This paper developed a scoring methodology and a map-based visualization tool to describe a supply chain's end of life impacts in three different areas: economic, environmental, and social. A key element in the scoring methodology is how it propagates uncertainty among heterogeneous data sources and borrows information from related data. Without other prior knowledge regarding the relation between data quality and

variance, the linear form for variances in Eqns. 1 & 2 is a reasonably simple and interpretable model since the results are only used for comparison purposes.

The tool compares available supply chain locations and lists specific sources by indicator for further research, with indicator weights that can be specified by the user. In the given case study, this weight sensitivity provided different and informative views of an example supply chain. Combined with the comparison map view, the gauges allow users to define priorities among the three impacts and explore possible alternative options. Further work could explore methods of visualizing or summarizing this sensitivity. ImpactMap was intended for use in the earlier process of a supply chain design and is rather a coarse-grained tool for directing more rigorous future investigation such as uses in final analysis or optimization, limited in other existing LCA software. The current development enables designers to explore potential supplier relationships or make product design decisions that will alter the material or manufacturing processes used without requiring the depth of modeling required by professional LCA tools.

**Acknowledgements.** The authors would like to recognize Ms. Yuchu Huang for her contributions to the project's data collection efforts, Dr. David A. Dornfeld and Dr. Margot Hutchins for helpful discussions, and The Social Hotspots Database for their contribution to the tool's social impact data. We also thank the anonymous reviewers for their insightful comments.

## References

- Cooper, J. S. and Kahn, E. (2012). "Commentary on issues in data quality analysis in life cycle assessment." *International Journal of Life Cycle Assessment*, 17 (1): pp. 499–503.
- Earthster (2012). Accessed from <http://www.earthster.org/> 06 May 2012.
- Economic Input-Output Life Cycle Assessment (EIO-LCA) (2008). US 1997 Industry Benchmark model. Carnegie Mellon University Green Design Institute. Accessed from <http://www.eiolca.net>. 06 May 2012.
- Hutchins, M. and Sutherland, J. (2008). "An exploration of measures of social sustainability and their application to supply chain decisions." *Journal of Cleaner Production*, 16 (15): pp. 1688-1698.
- Junnilla, S., and Horvath, A. (2003). "Life Cycle Environmental Effects of an Office Building." *Journal of Infrastructure Systems*, pp. 157-166.
- Lloyd, S. and Ries, R. (2007). "Characterizing, Propagating, and Analyzing Uncertainty in Life-Cycle Assessment: A Survey of Quantitative Approaches." *Journal of Industrial Ecology*, 11(1), pp. 161-179.
- Nagurney, A. and Nagurney, L. (2010). "Sustainable supply chain network design: a multicriteria perspective." *International Journal of Sustainable Engineering*, 3 (3): pp. 189-197.
- Neuman, S. P., Xue, L., Ye, M., and Lu, D. (2012). "Bayesian analysis of data-worth considering model and parameter uncertainties." *Advances in Water Resources*, 36 (1): pp. 75–85.
- Patagonia (2012). "The Footprint Chronicles: Our Supply Chain". Accessed from <http://www.patagonia.com/us/footprint/>. 06 May 2012.
- Sourcemap (2010). Accessed from <http://sourcemap.com/>. 06 May 2012.
- Social Hotspots (2010). "Social Hotspots Database." Accessed from <http://socialhotspot.org/>. 06 May 2012.
- Sustainable Minds (2012). "Eco-concept + LCA software." Accessed from <http://www.sustainableminds.com/software>. 06 May 2012.
- Wang, E. and Shen, Z. (2013). "A hybrid Data Quality Indicator and statistical method for improving uncertainty analysis in LCA of complex system - application to the whole-building embodied energy analysis." *Journal of Cleaner Production*, 43 (1): pp. 166-173.
- Yakovleva, N., Sarkis, J., and Sloan, T. (2011). "Sustainable benchmarking of supply chains: the case of the food industry." *International Journal of Production Research*, pp. 1-21.