The 6th International Conference on Life Cycle Management in Gothenburg 2013

STREAMLINED CARBON FOOTPRINT COMPUTATION — A CASE STUDY IN THE FOOD INDUSTRY

Yin Jin Lee and Edgar Blanco.*

** Massachusetts Institute of Technology, 77 Massachusetts Avenue, Building E40-261, Cambridge, MA 02139-4307. Email: yinjin@mit.edu.*

Keywords: carbon footprint calculation; uncertainty analysis; data gap; cut-off error

ABSTRACT

Using streamlined iterative carbon footprint calculation with uncertainty analysis, we developed a methodology to quickly estimate the carbon footprint of prepared meals. We used the methodology to calculate the carbon footprint of a dish that included 105 inputs from cradle to grave. We were able to quantify the uncertainty of data quality, data gap and cut-off error. We also used hypothetically true carbon footprints to validate the accuracy of the result. It was observed that the total carbon footprint estimate and its uncertainty converged when only the top 33 out of 105 inputs were updated. The method affirms that using cut-off strategies is a legitimate approach to reduce data collection effort while maintaining the integrity of carbon footprint estimates.

INTRODUCTION

Although the food service sector controls the types and the compositions of food made available to consumers, it has limited control and knowledge of the true carbon footprint of the various ingredients and processes. A methodology developed with a focus on calculating the carbon footprint of prepared food with minimal information can help food service providers design their dishes to minimize their carbon footprint and ideally make carbon labeling easier. For simplicity, we refer to all processes and materials that contribute to the total carbon footprint as *inputs*. In general, carbon footprint analyses can be greatly streamlined if we can focus our data collection effort on only a subset of all the inputs, with an understanding that these inputs are the ones that largely determine the final value of the carbon footprint.

In a previous case study (Lee, Yang, & Blanco, 2012), we noted several limitations in using existing data to compute food carbon footprint. These limitations are applicable to carbon footprint calculations in general too. Firstly, we cannot justify the subset size, commonly referred to as the *cut-off*. The British Standards Institute PAS2050:2011 recommends to cut off the inputs and outputs that contribute to the last 1% of the total carbon footprint (BSI British Standards, 2011) while the GHG Protocol recommends that the companies use screening calculations to prioritize the data collection effort (Bhatia et al., 2011). In this work, we will show that by performing CF and uncertainty calculation as the emission factor as the emission factors of the high impact inputs are corrected, one at a time, we can eventually

witness convergence of the carbon footprint estimate, as postulated by Williams (2009).

Secondly, we do not have a way to support the accuracy of the estimated carbon footprint when we use proxies to replace unknown emission factors. The lack of true emission factors is also called data gaps. This problem are generally addressed in two ways (Bhatia et al., 2011; Canals et al., 2011): first is by extrapolating from existing databases and second is by using existing emission factors as proxies in the calculation. Neither approach has been proved to be accurate while accounting for the differences between the known proxies and the unknown true emission factors. In this work, we would borrow the concept of *test sets* from data mining (Hastie, Tibshirani, & Friedman, 2009) to support the use of proxies for screening calculations, and as substitutes for the emission factors of the lower impact inputs.

Lastly, we introduce an analytical solution to find proxy variance when the proxy is the average of several standard emission factors. Weber (2012) calculated the variance of the *average proxy* assuming that the proxies have the same true mean. Yet in most applications, the emission factors do not share the same true mean. Assuming that the emission factors share the same true mean can underestimate the uncertainty of the average proxy. Patanavanich (2009) proposed using Monte Carlo simulation to obtain the correct variance for average proxy when the sources of emission factors are distinctly different. In this work, we found that the *Law of Total Variance* (Bertsekas & Tsitsiklis, 2002) is the analytical solution to the Monte Carlo simulations, and we applied it to simplify the uncertainty analysis.

METHODS

Primary data collection and emission factor database

The food production phases were divided into five distinct stages, namely, Agriculture, Transportation, Packaging, Preparation and Disposal. Primary data of the inputs to a chicken entrée was collected directly from a restaurant and its suppliers. There were 105 inputs to the dish. Separately, we consolidated a database of emission factors sourced from multiple sources. These data are available in the Lee (2013).

Test set

The proxy emission factors in the database were split into two sets, the first set was used for estimating the total carbon footprint and the second set was used as hypothetically true emission factors to test the accuracy of the estimate.

Analytical solution to proxy uncertainty

Here we describe how the *Law of Total Variance* is applied to find the variance of the average of multiple proxies. If the i^{th} input to the product has N available standard emission factors, each emission factor (EF_n) is assigned an uncertainty based on the analyst's judgment (Meinrenken et al., 2012). The total variance $var(\overline{EF_i})$ for the i^{th} input to the product is calculated in two parts, one part is the average of the N variances (represented by $\boldsymbol{E}[var(EF_n|N)]$), and the second part is the variance between the N averages (represented by $var(\boldsymbol{E}[EF_n|N])$). The complete representation of the equation is

$$var(\overline{EF_i}) = \boldsymbol{E}[var(EF_n|N)] + var(\boldsymbol{E}[EF_n|N]), for\ n = 1, \dots, N$$

where

$\overline{EF_i}$ is the average proxy for the i^{th} input obtained from all N emission factors EF_n
EF_n is the n^{th} proxy emission factor

Iterative approach
The iterative approach is as follows: a screening calculation is first done to compute the total carbon footprint of the product. The individual carbon footprint contribution from each input is ranked and the input with the greatest contribution to the total carbon footprint (also the input with the highest impact) is replaced by the hypothetically true emission factor in the test set. The total carbon footprint is recalculated with other inputs using the same proxies as in the previous step. We revise the total carbon footprint until all inputs are updated.

RESULTS AND DISCUSSION

Proxy uncertainty
The expected total carbon footprint and standard deviation of the *screening calculation* based on the analytical solution and the Monte Carlo simulation were 2.051±0.392 $kgCO_{2eq}$ and 2.043±0.383 $kgCO_{2eq}$ respectively. The results from the analytical solution and the results from the analytical solutions matched even though they calculated the variance with different approaches. While the simulation results give the distribution of the carbon footprint estimate, the analytical solutions can substantially cut down the computing time and the amount of data generated in the intermediate steps.

Using iterative approach to justify the cut-off
Based on the blue curve in Figure 1, the carbon footprint estimate was 2.34±0.06 $kgCO_{2eq}$ when the top 33 inputs (99^{th} percentile) were updated. The final estimate when all 105 APs were updated with the test set emission factors was also 2.34±0.06 $kgCO_{2eq}$. We can use this graph to choose other cut-offs depending on our need for accuracy and precision. By updating the total carbon footprint step-wise with test set(s), analysts can now justify the way they choose cut-off. This approach can be applied to carbon footprint calculations in other industries too.

Test set
The significance of using hypothetically true emission factors in a test set showed that the carbon footprint estimate would converge quickly even when only a subset of the inputs have the hypothetically true emission factors, while the remaining inputs are represented by proxies (Figure 1). This is possible because the inputs were updated in descending order of their carbon footprint impact, and because a small subset of the inputs contributed to a large fraction of the total carbon footprint. This trend supports that appropriate cut-offs can save substantial data collection effort without sacrificing the accuracy of the estimate.

CONCLUSIONS

Using food carbon footprint calculation as a case study, we found several ways to make carbon footprint calculations easier and more accurate. First is to use multiple proxies, and to calculate their uncertainty using the Law of Total Variance. We agree with Meinrenken et al. that analytical methods for error propagation can make calculation and analysis easier and that analytical methods for error propagation can make calculations faster (Meinrenken et al.,

2012). Next we showed that by recalculating the total carbon footprint and its uncertainty as we updated the high impact inputs one at a time with hypothetically true emission factors, we could quantitatively justify cut-off, which at this point, is still an arbitrary assigned standard.

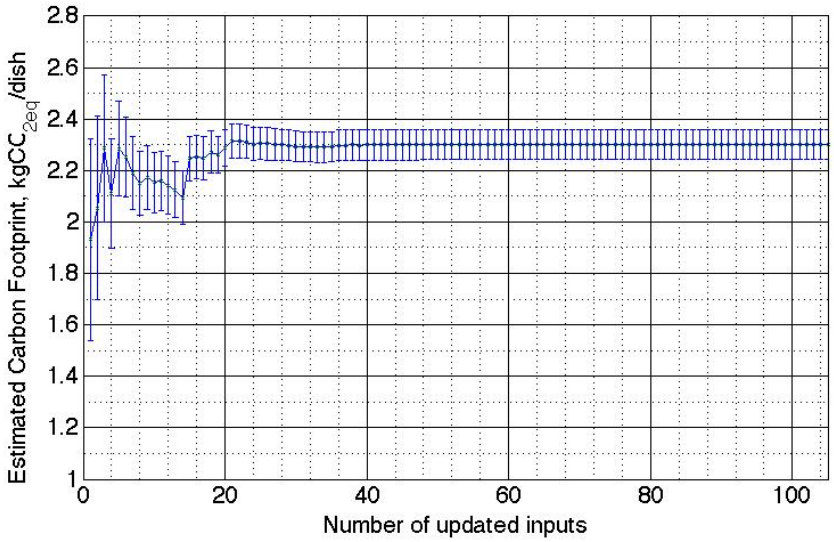

Figure 1. The estimated carbon footprint estimate converges quickly as the inputs are updated with the hypothetically true values from the test set.

REFERENCES

Bertsekas, D., & Tsitsiklis, J. (2002). General Random Variable. *Introduction to Probability* (2nd ed., p. 192). Belmont, MA: Athena Scientific.

Bhatia, P., Cummis, C., Brown, A., Draucker, L., Rich, D., & Holly Lahd. (2011). *GHG Protocol: Product Life Cycle Accounting and Reporting Standard*. USA: World Resources Institute. Retrieved from http://www.ghgprotocol.org/standards/product-standard

BSI British Standards. (2011). *PAS 2050:2011 Specification for the Assessment of the Life Cycle Greenhouse Gas Emissions of Goods and Services*.

Canals, L. M. I., Azapagic, A., Doka, G., Jefferies, D., King, H., Mutel, C., Nemecek, T., et al. (2011). Approaches for Addressing Life Cycle Assessment Data Gaps for Bio-based Products. *Journal of Industrial Ecology*, *15*(5), 707–725. doi:10.1111/j.1530-9290.2011.00369.x

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd Editio.). New York, NY: Springer.

Lee, Y. J., Yang, X., & Blanco, E. (2012). Streamlined life cycle assessment of carbon footprint of a tourist food menu using probabilistic underspecification methodology. *Sustainable Systems and Technology (ISSST)*. Boston, MA: IEEE International Symposium. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6227980

Meinrenken, C., Kaufman, S., Ramesh, S., & Lackner, K. (2012). Fast Carbon Footprinting for Large Product Portfolios. *Journal of Industrial Ecology*, *00*(00), 1– 11. doi:10.1111/j.1530-9290.2012.00463.x

Weber, C. L. (2012). Uncertainty and Variability in Product Carbon Footprinting. *Journal of Industrial Ecology*, *16*(2), 203–211. doi:10.1111/j.1530-9290.2011.00407.x