Life Cycle Emissions Distributions within the Economy: Implications for Life Cycle Impact Assessment

Short Title:
Life Cycle Emission Distributions

Gregory A. Norris
Sylvatica, 147 Bauneg Hill Road, North Berwick, Maine, 03906, and
Harvard School of Public Health, Boston, MA

Presentation for Society for Risk Analysis 2000 Annual Meeting, Washington D.C.

2000.12.03-06

Key Words: Life Cycle Assessment, Life Cycle Inventory Analysis, uncertainty analysis, boundary truncation, Input/Output Analysis, Impact Assessment, streamlining

Abstract

Refinements of methods for life cycle impact assessment are directed at reducing errors and quantifying and reducing uncertainties in results. The uncertainty reduction benefits of such refinements depend upon the structure of the inventory model. The structure of inventory models in general are investigated using an economic input/output life cycle assessment model of the US economy. Percentiles for the share of total upstream emissions contributed by the set of processes in each supply tier are presented for US commodities and several important pollutants. Capturing at least 90% of the total direct plus upstream emissions for criteria air pollutants and toxic releases for at least 75% of the commodities in the US economy requires full modeling of direct emissions plus the first five supply tiers.

A method is developed and applied for streamlining Input/Output models. The method yields several conclusions relevant to risk-based life cycle impact assessment. The depth and breadth requirements for capturing a high percentage (e.g., > 80%) of total emissions vary widely across products or commodities. Models focusing on just direct plus the 15-20 most important processes in tier 1 can capture a median of just over 40% of emissions. To capture more than 60% of total emissions for more than half of all commodities requires models requires models with more than 4000 process instances. To well characterize the total impacts of products, life cycle impact assessment methods must characterize foreground process impacts in a site-informed way and mean impacts of far-removed processes in an unbiased way.

Introduction: Source properties influence impacts of a release

LCA researchers are developing and refining methods for impact assessment that integrate a variety of approaches to fate and exposure modeling as well as (in some instances) effects modeling. This is true for a variety of impact categories. In each case, the modeling is used to derive a set of "characterization factors" which estimate the relative expected strength of influence, within a given impact category, of equivalent mass quantities of pollutant release, *as a function of pollutant species*. [e.g., Udo de Haes et al., 1999]. In some cases, the characterization factors are also allowed to vary as a function of *the initial receiving media* [e.g., Hertwich et al. 2001; Goodkoop and Spriensma 2000].

More recently, regionalized factors have been developed, which take into account the influence of source location on expected impacts [Potting et al. 1998; Nishioka et al. 2000]. These advancements have shown that the location of a release can have a strong influence upon its expected impact, at least for some impact categories and some different sets of locations. But the emissions in a life cycle occur at many different locations, not just one. Does site-specificity really reduce uncertainty in life cycle assessment results?

In order to take location into account in life cycle impact assessment, we require information about the locations of the processes, which is a non-trivial information requirement. Before we can assess the need for, and uncertainty or error reductions that result from, regionalized characterization analysis in life cycle impact assessment, we need quantitative answers to the following questions:

a) From how many sites, in what percentages, do life cycle emissions originate?
b) What geographic dispersion and/or autocorrelation is exhibited in the set of emissions sites?
c) How do these results depend on the class of pollutant?
d) How do they depend on the class of product or process whose life cycle is being considered?

For many processes in a life cycle supply chain, their locations cannot be known with certainty, only estimated. Two additional questions for regionalized LCIA are therefore:

e) With what certainty can process locations be specified?
f) Given the uncertainty in location specification, how much uncertainty reduction in total final LCIA results is yielded by regionalized Life Cycle Impact Assessment?

Herein we present initial answers to questions a, c, and d, from a larger research effort which is motivated by all of the above questions.

Approach

Empirical investigations into this subject have been undertaken using input/output LCA (IO LCA) models of the US economy. The research has been undertaken with a 500-sector IO LCA model for the US that has been constructed from databases published by the US government. Databases from the US Department of Commerce describe the flows of goods and services among the sectors in monetary terms (Kuhbach and Planting 2001; Lawson 1977). These can be used to estimate tier-by-tier the economic activity in the supply chain for each of 500 commodity groups. Together these 500 commodities span the entire spectrum of commodities bought and sold in the US.

Separate databases from the US EPA report annual pollution releases from each sector. Here we make use of the emission trends inventory (e.g., EPA 2000) which reports emissions of criteria air pollutants from point sources by industry category. We also used estimates of sector-specific fossil fuel combustion, by fuel type, derived from the Manufacturing Energy Consumption Survey of the US Department of Energy (e.g., EIA 1997) in order to estimate sector-specific emissions of fossil-based CO2. Finally, we used data on toxic releases from reporting establishments within reporting sectors from the US EPA's Toxic Release Inventory (TRI) (see, for example, EPA 2001) to obtain sector-level sums of toxic releases by receiving media. These data on annual sector total releases for each pollutant type are then divided by the annual economic output from each sector to derive annual average pollution coefficients for each sector. The resulting coefficients are used with the supply chain computations to estimate supply chain pollution upstream of each commodity.

Results

Tier-wise convergence to upstream totals

The concept of supply tier is straightforward. The set of all suppliers of the direct inputs to a given using sector are termed that using sector's "first tier suppliers." The set of all the direct suppliers to these first-tier suppliers comprise the second supply tier of the original using sector, and so-on. Tier "zero" is the final sector producing the commodity itself.

First, we computed "percentiles" for the cumulative upstream pollution by tier for the full set of US commodities, for each of the US EPA's "criteria air pollutants" (NOx, VOCs, particulates, CO, and SO2), for CO2, as well as for toxic releases to air, water, land, and underground as reported in the US Environmental Protection Agency's Toxic Release Inventory (TRI). The tier-wise cumulative percentiles indicate the share of total commodities for which cumulative emissions up to and including that tier in their upstream life cycle account for less than that percentage of the commodity's total upstream emissions for that pollutant. Thus, the percentiles can be used to judge the probability that for some randomly chosen commodity, modeling a given number of tiers will capture a given fraction of total upstream emissions for a given pollutant.

As an example, Figure 1 presents the cumulative percentiles for emissions of sulfur dioxide, SO2. The curve for the 5^{th} percentile in Figure 1 indicates for only 95 percent (100% – 5%) of commodities (goods and services) produced in the US economy, upstream models which include the 0^{th} through 3^{rd} tiers will capture at least 75% of the total upstream emissions of SO2. The figure also indicates, for example, that for 75% of the commodities, models which span tiers 0 through 4 will capture at least 90% of the total upstream emissions of SO2.

Figure 2 presents similar results for releases of volatile organic compounds (VOCs). The differences between Figures 1 and 2 illustrate that the *speed of convergence* – that is, the probability or percent of commodities for which a given number of tiers captures the bulk of the upstream emissions – this speed of convergence varies from pollutant to pollutant. These differences are captured and summarized in Figure 3, which presents the 25^{th} percentiles for all of the pollutants or pollutant categories in the I/O LCA modeling system. These curves indicate that toxic releases to water and to land are the slowest pollutants to converge, with 25% of all commodities still missing approximately 40% of their total upstream emissions after tiers 0 through 3 have been modeled. At the other end of the spectrum are CO2, toxic releases to air, and a measure of total manufacturers' waste treatment and disposal costs ("Waste T&D") all converge more rapidly, with tiers 0 through 3 accounting for over 80% of the upstream total for at least 75% of the commodities. A summary conclusion is that to capture at least 90% of the total direct plus upstream emissions for criteria air pollutants and toxic releases for at least 75% of the commodities in the US economy requires full modeling of direct emissions plus the first five supply tiers.

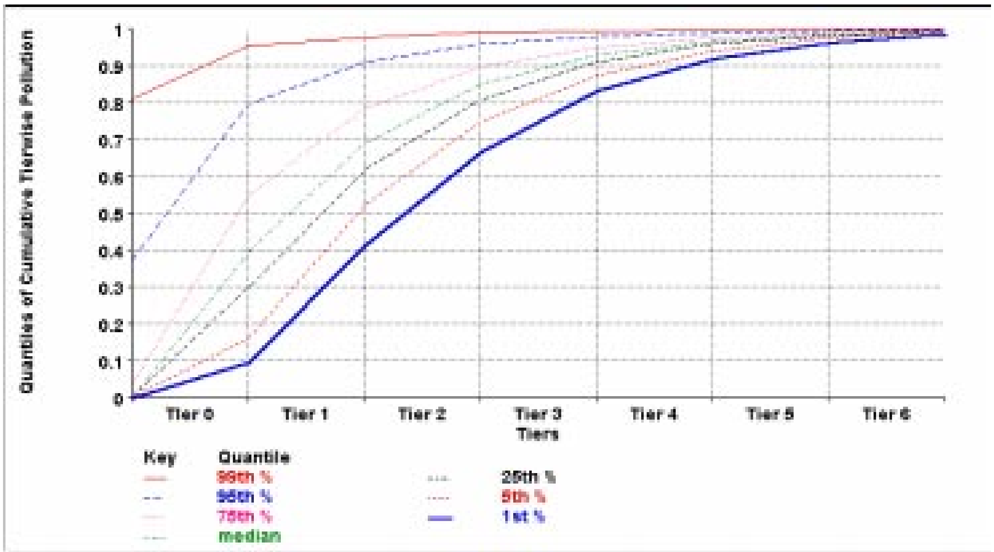

Figure 1: Percentiles for Cumulative Upstream Releases of SO2

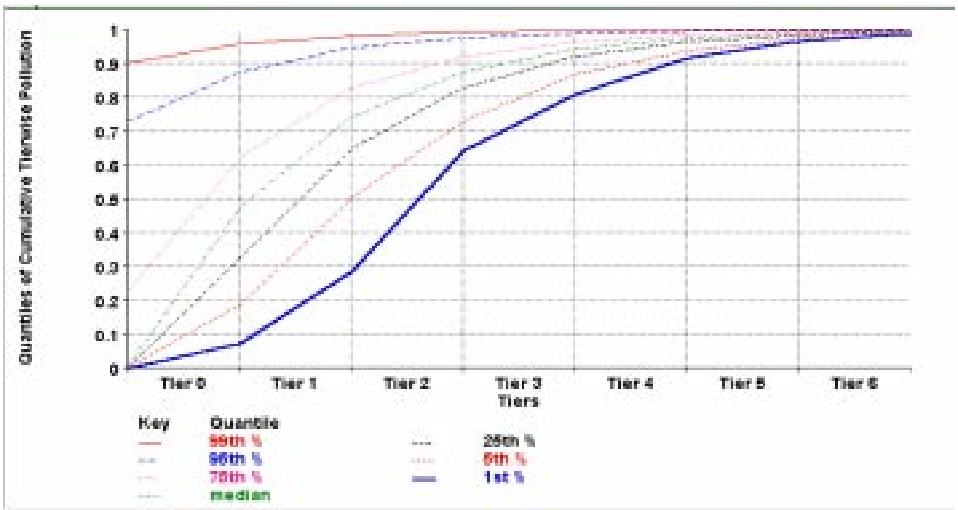

Figure 2: Percentiles for Cumulative Upstream Releases of VOCs

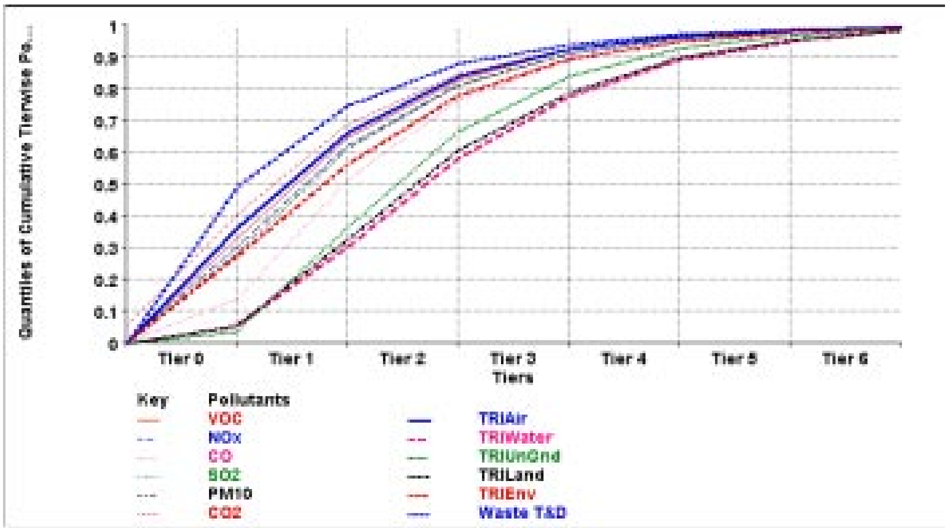

Figure 3: 25th Percentiles for Upstream convergence, by pollutant type; convergence is slower than indicated for 25% of the commodities in the US economy

Contributions within Tier 1

Next, we look within the tier 1 to estimate percentiles for the numbers of individual sites contributing the bulk of the emissions for each pollutant for this first tier in the supply chains of each originating commodity. These results are summarized in Figures 4 and 5. Figure 4 presents the results for a specific pollutant (CO2), while Figure 5 presents the results for the 10th percentile, or 90 percent of commodities. Figure 5 indicates that depending upon which pollutant is selected, capturing 90% of the first tier's emissions for 90% of the commodities requires modeling only the top 5 polluters in the tier for SO2, while it requires modeling the top 23 emission sources in the tier for toxic releases to air.

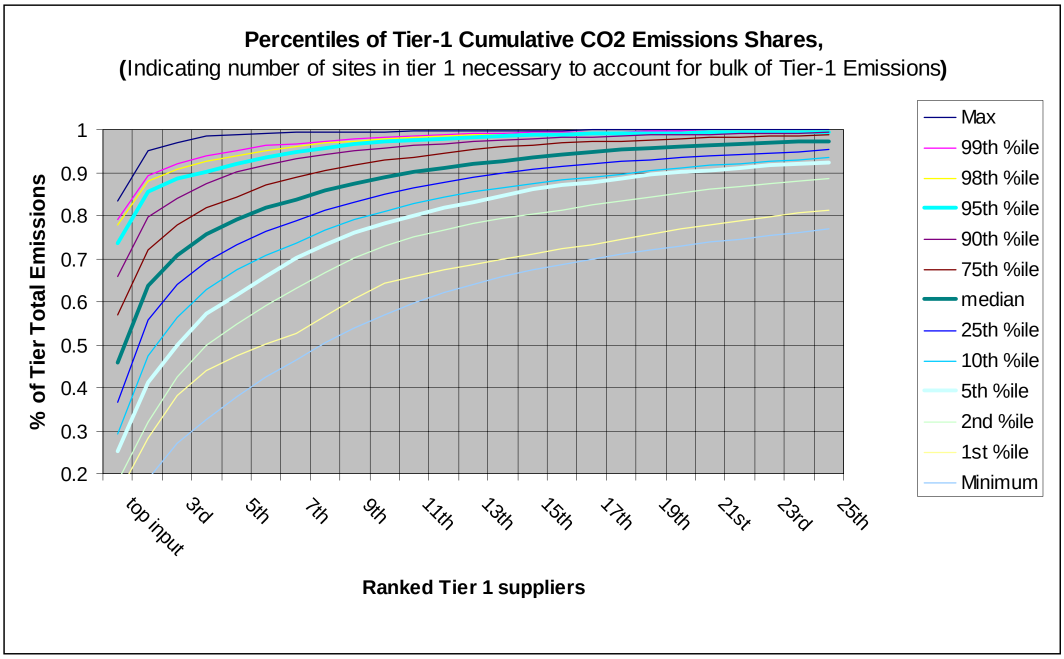

Figure 4: Percentiles of Tier-1 Cumulative CO2 Emissions Shares

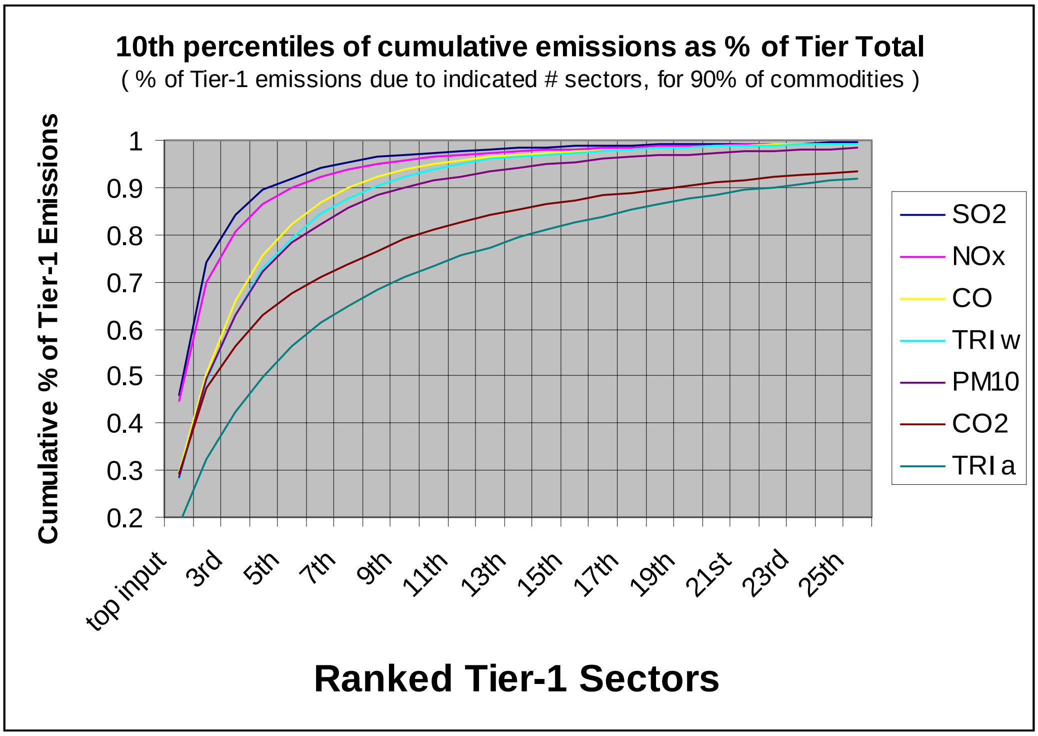

Figure 5: Tenth Percentiles of Cumulative Emissions as % of Tier Total

Contributions within the Entire Supply Chain: Process Instances

First we address the question of how many sites in the supply chain make a *non-zero* contribution to total upstream emissions. As Figure 3 made clear, inclusion of the direct emissions plus those from the first 6 tiers is necessary to capture at least 95% of the upstream emissions reliably across pollutant types. If we neglect sites beyond the 6^{th} tier, how many sites are there in this 6-tier model?

Most processes in an LCA supply chain model are characterized using data intended to be representative of their process *type*, (e.g., rolling of steel sheet, coal steam production of electricity in the US, transport by Dutch rail, average plywood production in the Southeast US) rather than by reference to a specific factory or process somewhere in the world. In a comprehensive model, processes of a given type appear multiple times. For example, many processes in the supply chain may use as an input steel manufactured in North America. Each "instance" of this steel use is associated with at least one instance of steel production, the location of which would need to be known in a purely site-specific LCA.

We define the number of process *instances* as the number of times within a supply chain model that process identities would need to be specified in order to use site-specific information for each process in the chain. Note that the number of steel plant instances in a very comprehensive supply chain model may exceed the actual number of steel plants in the world. Thus, each instance of the process type "steel mill" in a supply chain model is not *necessarily* a different site; instead, each instance is another usage of steel for which the identity of the supplier could in principle be specified.

The concept of process instance bears upon the proper design of sampling procedures for uncertainty analysis in life cycle analysis. Suppose, for example, that a frequency distribution can been specified which characterizes variability in the expected exposure from a kg of particulate emissions from steel mills. Further suppose there are 22 instances of steel mill in the supply chain, each (for simplicity here) expected to release equal quantities of particulates per functional unit. In this simple example, Monte Carlo estimation of a probability density function for the expected total exposure would be generated by sampling 22 times, with replacement, from the frequency distribution for expected exposure. Thus, the number of instances of important processes in the supply chain governs the uncertainty in estimates of their impact, and the uncertainty reductions brought by site-specific characterization factors.

An ordered plot of the numbers of suppliers to sectors in the baseline Input/Output model is shown in Figure 6. From this plot we see that the number of inputs to sectors ranges from a low of 144 to a maximum of 432 in the 475-sector model. For 90% of the sectors, this number of inputs ranges between 300 and 405, and the median number of inputs is 375. Based on this median number of inputs, an approximation for the total number of instances sites in the first 6 tiers 375^6 which exceeds 2.7 quadrillion ($2*10^{15}$).

The number of process instances in a 6-tier model is enormous, but how many sites in the supply chain make a *significant* contribution to total upstream emissions? Theoretically one could calculate the instance-by-instance emissions for all 3 quadrillion of the process instances in tiers 1 through 6, then sort them and rank them. However, this calculation is neither practical nor necessary. We seek upper and lower bounds on the number of important instances in a 6-tier model. Three quadrillion provides a high upper bound. A more practical approach to finding a lower bound is to study the emissions for a streamlined or "pruned" process tree model that has includes only the most important upstream linkages.

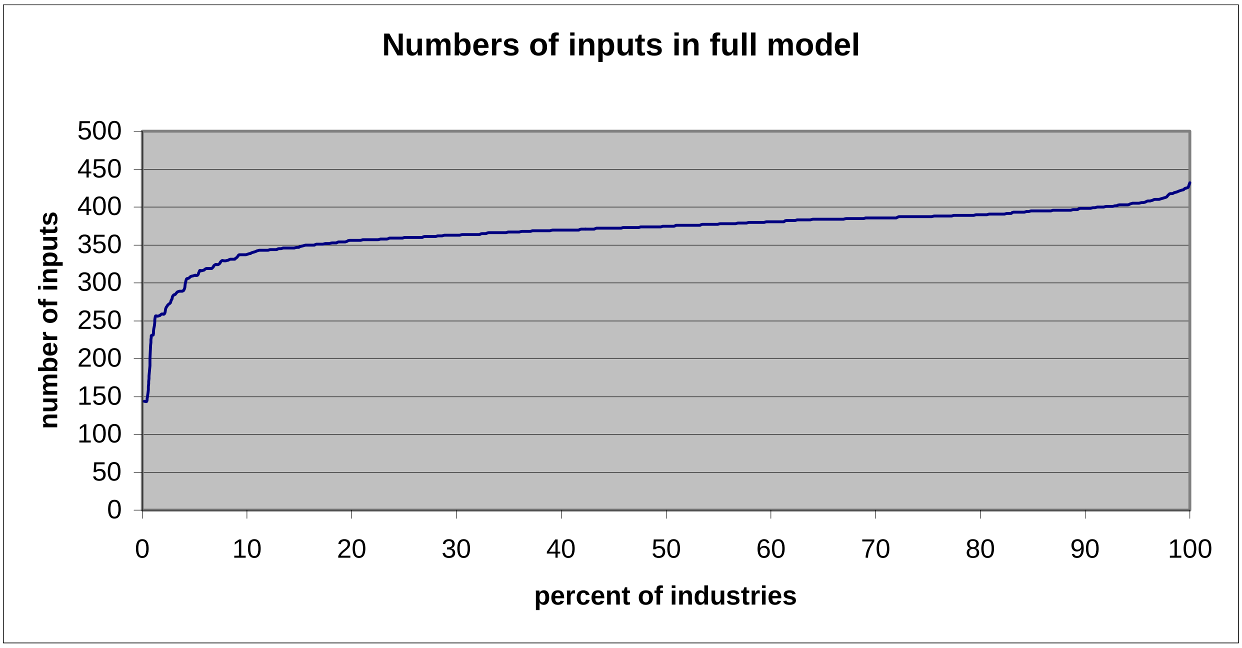

Figure 6: Number of inputs to industries in initial 500-sector EIO-LCA model of the US economy

A streamlining procedure has been created and applied which retains in the I/O LCA model only the minimum set of inputs to each sector in order to account for a specified fraction of the total upstream releases of a given emission. For each of the 475 industries in the input/output LCA model the total upstream releases of a given emission associated with each input to that industry are calculated; here we present the results for a first application of the method to fossil CO2 emissions. The inputs to each industry are then ranked in terms of their CO2 emissions upstream of that industry. Three separate streamlined models are developed that retain only the minimum set of inputs to each industry required to capture 90%, 95% or 99% of direct plus upstream releases.

Figure 7 shows that considerable reductions in the number of inputs per sector are possible, even at the 99% threshold. A median of 90 inputs is retained when the threshold is set to 99%. A median of 30 inputs is retained in the 95% threshold model. For the 90% threshold, all sectors retain fewer than 40 inputs, and a median of only 16 inputs are retained.

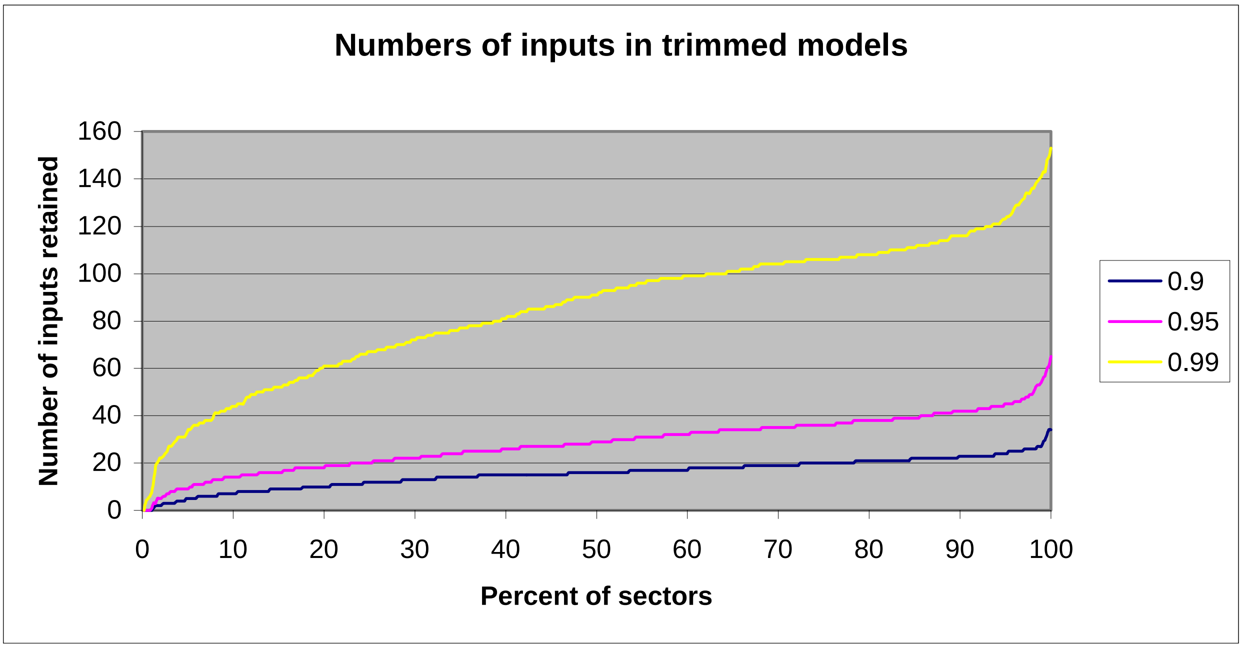

Figure 7: Input retention in the streamlined I/O-LCA models

The performance of the streamlined models is summarized in Figures 8, 9 and 10. Figure 8 illustrates the compounding effects of streamlining upon total cumulative results. That is, while each sector has been pruned to include enough inputs to capture 95% of its direct plus upstream burden (assuming that each input is modeled with 100% thoroughness), now each of these inputs is also modeled in a streamlined way, and so-on up the supply chain. We see that the 95% threshold model captures a median of 77% of total (direct plus 6 tiers) CO2 emissions overall. This median is calculated over the full set of 475 commodities in the economy. For the 90%-based model, the median cumulative coverage is about 65%.

Note also in Figure 8 that the more streamlined a model is, the more quickly it approaches its asymptotic total upstream emissions value. While tier 6 emissions still add perceptibly to the total upstream burden for the full model, we can see that by tier 3 the 90% streamlined model has essentially reached it's asymptotic value. In fact, the penalties of streamlining are higher for more distant tiers. This can be seen best in Figure 9, which displays median results for the share of full-model tier-specific CO2 emissions that is captured by each of the streamlined models. For example, the 90% model captures less than half of the total tier 2 emissions accounted for by the full model. By tier 4, this 90% streamlined model is able to account for only 10% of the tier-4 emissions reported by the full model. In other words, the processes contained in the streamlined model account for smaller and smaller shares of each tier's total emissions, the farther out in the supply chain we look.

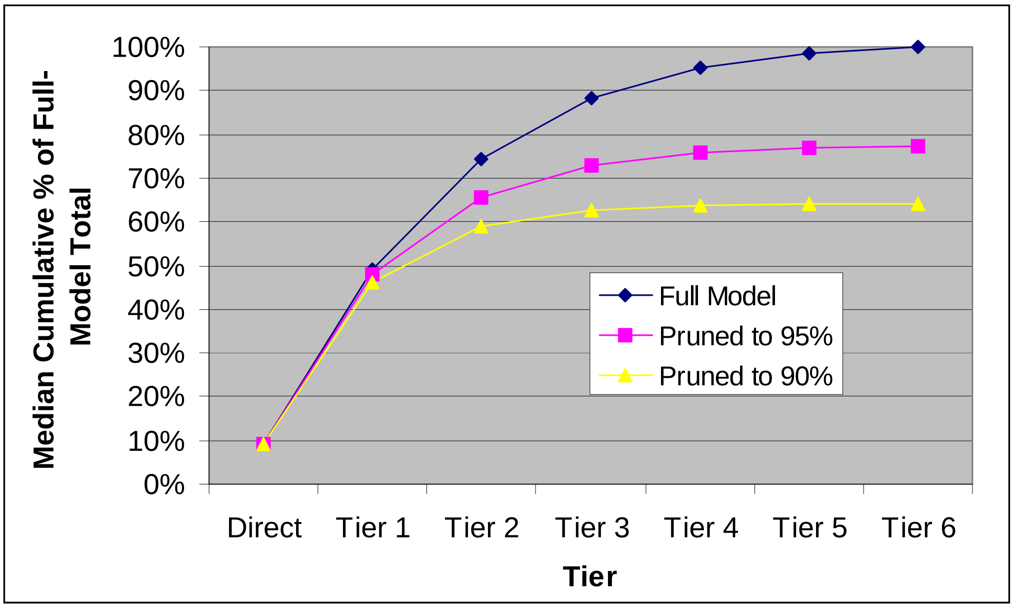

Figure 8: Median cumulative tier-wise percentage of full-model 6-tier direct plus upstream CO2 emissions

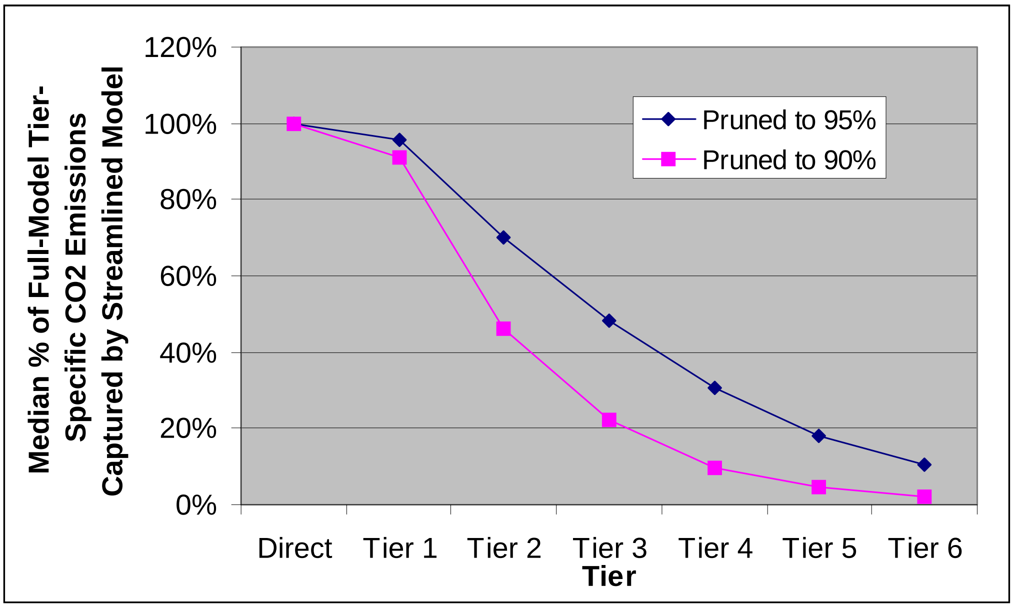

Figure 9: Median results for share of full-model tier-specific CO2 emissions captured by streamlined models

The share of total upstream emissions captured by the streamlined models varies considerably across products, even though all sectors of the economy have been streamlined to the same threshold level. Figure 10 provides a plot of percentiles (across the 475 products) for the share of total (Tier 0 through 6) emissions accounted for by the cumulative results of the 95%-streamlined models. For 5% of products, tier 0 plus the 16 sites in tier 1 account together for at least 80% of the total emissions. The total emissions coverage provided by the streamlined model (by tier 6) ranges from 63% to 94%. We conclude that products vary in terms of the total share of their emissions accounted for by the "main line" core of key process linkages retained in the streamlined model.

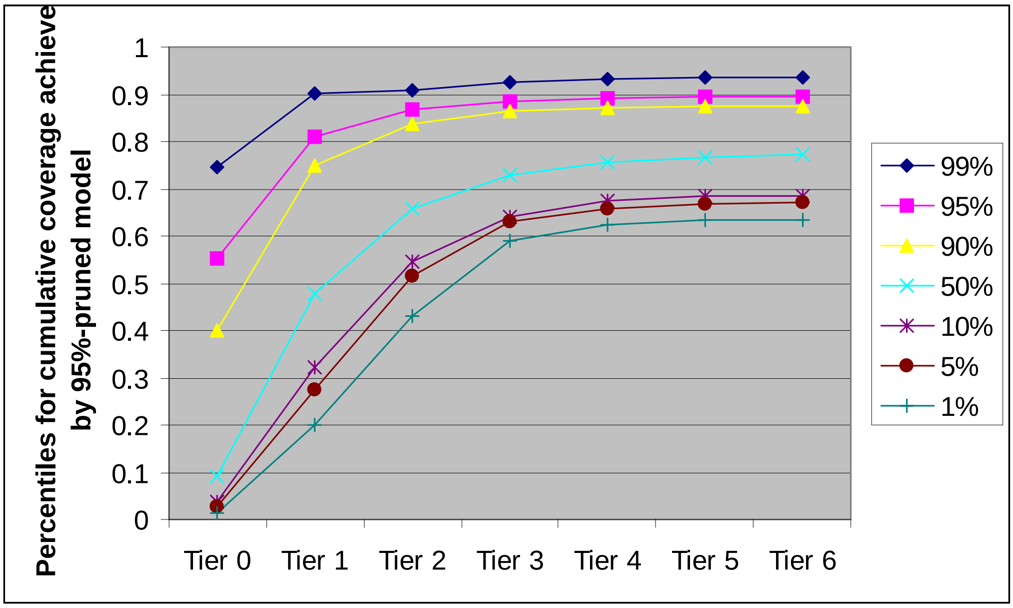

Figure 10: Variation in cumulative coverage (by tier) of total direct plus upstream CO2 provided by the 95%-threshold streamlined model

Results from these streamlined models provide information about the minimum number of process instances required to achieve selected levels of overall emissions coverage. Recall from Figure 7 that the 95% streamlined model retained a median of 30 inputs. In this model we expect a median of approximately 900 process instances in tier 2, and 27,000 in tier 3. However, Figure 10 shows us that the median coverage provided by this model through tier 3 was just over 70% of total emissions. For the 90% model, an estimated median of 4250 instances through tier 3 accounts for just over 60% of total emissions. Thus, we can state in general that except for the small fraction (e.g., 10%) of products represented in the upper left portion of Figure 10, even one thousand process instances cannot be counted on to provide more than 50% of emissions coverage for the typical product.

Conclusions

The results presented here are tentative for several reasons. First and foremost, we conducted model streamlining and assessment based on just one pollutant. Second, the economic data in the model reflect the 1992 US economy, and the pollutant emissions data correspond to the mid-1990s. Third, while CO2 emissions coverage in the model is quite complete, most other emissions inventories are less incomplete, due to factors such as underreporting within the TRI reporting sectors, non-reporting of toxic releases by sectors outside the TRI's scope, and inability to assign all of the EPA's area and mobile source emissions to specific economic sectors. Still, the results provide initial findings of importance.

First, the depth and breadth requirements for capturing a high percentage (e.g., > 80%) of total emissions vary widely across products or commodities; some can be well-modeled with a few early-tier processes, while others require very broad and deep models. Development of a streamlined I/O-LCA model showed that a median of just over 40% of emissions can be captured by models focusing on just direct plus the 15-20 most important processes in tier 1. To capture more than 60% of total emissions for more than half of all commodities requires models requires models with more than 4000 process instances.

For life cycle impact assessment, this means suggests a two-part or combined approach. High importance (direct, first- and perhaps a few second-tier processes) make a large enough contribution for some products to warrant a site-informed or even site-specific approach to impact assessment. At the same time, the results also show that far-removed processes (in supply tiers 3-6) can account for roughly half of total supply chain emissions for many products. Thus, accurate estimates of total supply chain impacts will require unbiased estimates of the mean impacts per unit emission for processes throughout the economy.

The emissions capture of streamlined (pruned-breadth) models falls sharply for outer tiers. This means that while pruning has little effect on estimates for tier-1 and tier-2 emissions, it has a very large impact on outer-tier estimates. Evidently, most outer-tier releases come from a very large number of very small contributions by literally billions of sites, rather than a few thousand "mainline" processes. A corollary to this finding is that streamlined models systematically underestimate the importance of model depth on capturing the full emissions consequences of products.

References

Kuhbach, Peter D. and Mark A. Planting 2001. "Annual Input-Output Accounts of the US Economy, 1997. *Survey of Current Business, 81:* 9-43.

Lawson, Ann M. 1997. "Benchmark Input-Output accounts for the US Economy, 1992: Make, Use and Supplementary Tables," *Survey of Current Business, 77:* 36-82.

US Environmental Protection Agency 2000: *National Air Quality and Emissions Trends Report, 1998.* EPA 454/R-00-003, Office of Air Quality Planning and Standards, Research Triangle Park, NC. March.

US Environmental Protection Agency 2001: Toxic Release Inventory Homepage Website: http://www.epa.gov/tri/.

Energy Information Administration 1997. *Manufacturing Energy Consumption Survey, 1994.* DOE/EIA-0512(94). Washington, DC. December.

Udo de Haes, H.A. (chairman), O. Jolliet (vice-chairman), G. Finnveden, M. Hauschild, W. Krewitt, R. Mueller-Wenk, 1999. "Best Available Practice Regarding Impact Categories and Category Indicators in Life Cycle Impact Assessment." *International Journal of Life Cycle Assessment* 4 (2): 66-75.

EG Hertwich, SF Mateles, WS Pease, TE McKone, 2001. "Human Toxicity Potentials for Life Cycle Assessment and Toxics Release Inventory Risk Screening." *Environmental Toxicology and Chemistry* **20** (4):928–939.

Goedkoop, M. and R. Spriensma, 2000. *The Ecoindicator 99: Methodology Report, Second Edition.* PRe Consultants, Amersfoort, Netherlands. April.

Potting, J., W. Schopp, K. Blok, M. Hauschild, 1998. "Comparison of the acidifying impact from emissions with different regional origin in life-cycle assessment." *Journal of Hazardous Materials* 61: 155-162.

Nishioka Y., J. Levy, P. Hofstetter, G. Norris, A. Wilson, J. Spengler, 2000: "Regional health effects of demand-side energy management using exposure efficiency," *Proceedings, 4th International Conference on EcoBalance,* Tsukuba, Japan.