UC Berkeley

UC Berkeley Previously Published Works

Title

A conceptual framework for uncertainty-informed system boundary expansion in life cycle assessment

Permalink

https://escholarship.org/uc/item/7d04z53d

Journal

Journal of Cleaner Production, 577

ISSN

0959-6526

Authors

Dai, Tao

Scown, Corinne D

Publication Date

2026-10-01

DOI

10.1016/j.jclepro.2026.149453

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike License, available at https://creativecommons.org/licenses/by-nc-sa/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

A Conceptual Framework for Uncertainty-Informed System Boundary Expansion in Life Cycle Assessment

Tao Dai[1,2], Corinne D. Scown[1,2,3]*

[1]Life-Cycle, Economics and Agronomy Division, Joint BioEnergy Institute; Emeryville, CA 94608, USA.

[2]Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[3]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[*]Correspondence: taodai@lbl.gov

Abstract

Life cycle assessment (LCA) is the most comprehensive framework for evaluating the environmental impacts of products and services, yet defining system boundaries remains a subjective exercise and can lead to unknown truncation error and complications in comparing across studies. Here, we introduce an uncertainty-informed, iterative boundary-expansion algorithm that quantitatively guides boundary expansion by answering which processes in the foreground system should be included and expanded to upstream detail and which may be aggregated or omitted. The proposed algorithm starts with the foreground system and identifies hotspot processes by comparing predefined thresholds to a process's mean and upper-tail contribution share, computed from Monte Carlo results that propagate uncertainty in both a process's input amount and emission intensity. We developed a stochastic simulation framework and created synthetic production system trees to examine the behavior of the proposed algorithm under controlled conditions. Our results show that the cutoff threshold is the dominant factor in determining final model complexity, while uncertainty-driven expansion is most beneficial in low-data-quality scenarios. Foreground impact concentration could shape the distribution of effort required for system boundary expansion, suggesting that conventional LCA practices focused on primary supply chains may systematically miss significant impacts in diffuse systems. Applying the proposed approach could potentially enhance the documentation of system expansion and improve

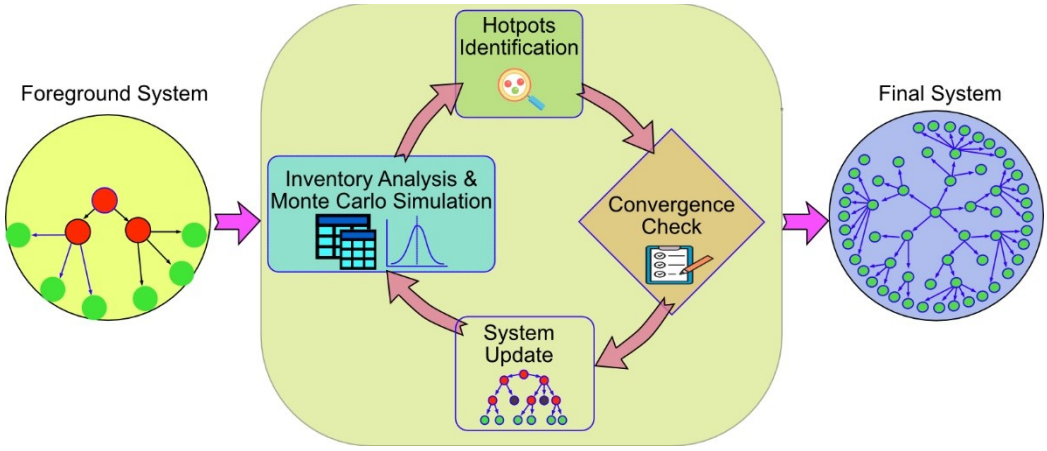
uncertainty analysis within LCA, ultimately increasing reproducibility and transparency in environmental assessments.

Keywords: Life Cycle Assessment; System Boundary Definition; System Boundary Expansion; Environmental Impact Assessment; Uncertainty Quantification; Product System

Highlights

- A framework for uncertainty-informed LCA system boundary expansion was proposed
- A stochastic simulation framework was developed for testing LCA under controlled conditions
- Considering both impact significance and data confidence could help mitigate truncation error

Graphic Abstract



1. Introduction

Life cycle assessment (LCA) plays a vital role in environmental decision-making (EC-JRC, 2012; Garcia and Freire, 2014). This ISO-standardized approach examines a product or service from a "cradle-to-grave" perspective, aiming to quantify its cumulative environmental impact across all life cycle stages (ISO, 2006a, 2006b). This comprehensive scope, however, also introduces practical time, resource, and data availability constraints, leading to unavoidable simplifications that omit elements of the system (often referred to as truncating or truncation error) across all LCAs (Bala et al., 2010; Douziech et al., 2021; Ross and Cheah, 2018; Scrucca et al., 2020; Zargar et al., 2022). LCAs typically start with foreground system analysis, where primary data are collected, and then expand to broader background systems, which are modeled using secondary data. The ISO standards provide limited quantitative guidance beyond the use of cut-off criteria (i.e., rules for omission justified in relation to the magnitude of contributions), which leads to the well-recognized paradox in LCA: practitioners would have to fully assess the excluded processes to justify excluding them (Clark et al., 2022; Dai et al., 2020a). Current LCAs often predefine system boundaries around the main supply chain and treat other upstream branches as negligible or adequately represented by emission factors. However, aggregating high-impact branches could potentially lead to hotspot omission or misattribution, as the representativeness of such a system relies heavily on the assumption that the main supply chain dominates the overall impact. The selection of main flows can also be biased toward one or a subset of metrics, such as CO_2-equivalent emissions, thereby missing supply chain elements that dominate other metrics in their scope, such as water use, acidification, or eutrophication. We therefore argue that the central challenge in achieving quantitative system boundary identification lies in moving beyond the predefined supply-chain approach and developing a systematic procedure for determining which processes in the foreground system should be included and expanded and which can be omitted.

Early efforts to address the challenge of system boundary definition (e.g., Tillman et al., 1994) recognized that a product system is inherently tree-shaped, with the functional unit acting as the root and the hotspots of material and mass flows forming the branches that expand in a stepwise manner. This analogy has become the standard approach in attributional LCA due to its simplicity and applicability, although Suh et al. (2004) demonstrated that the truncation error introduced by

the cutoff of small impact flows can be significant, as it ignores the economy's infinite, circular nature, and suggested using a hybrid approach that integrates input-output LCAs to address this. The hybrid approach, while conceptually powerful, has only been adopted by a small fraction of LCA studies (e.g., Scown et al., 2011; Yue et al., 2016), hindered by data scarcity, complexity, and the introduction of aggregation errors. Consequently, researchers revisited the tree analogy through the implementation of computational search methods, such as the depth-first search algorithms in Li et al. (2014), to facilitate quantitative boundary identification and move beyond a predefined system boundary, which, nevertheless, could still remain limited by their deterministic nature, reliance on data completeness, and the conflation of mass flow with environmental importance. Despite these advances, boundary identification still largely remains deterministic, so that the uncertainty in inventories and characterization factors is rarely propagated into the system boundary determination processes. In addition, there is no standardized platform or public benchmark datasets for systematically evaluating the consequences of system boundary determination; as a result, existing studies often rely on isolated case studies, limiting replicability and generalization of findings across product systems.

Advances in inventory streamlining and uncertainty quantification in LCA have laid the technical groundwork for uncertainty-informed system boundary identification. For example, Olivetti et al. (2013) proposed a structured under-specification approach that utilizes aggregated categorical-level impact information to estimate emission factors and the associated probabilistic distributions for missing data. This approach shows promise for application to a variety of material categories and can be further improved to fully quantify uncertainty (Dai et al., 2022, 2020a; Rodrigues et al., 2018; Tecchio et al., 2019). Furthermore, the increasing application of machine learning approaches (e.g., random forest and large language models) in inventory streamlining can enhance data availability and uncertainty quantification by predicting multiple values and incorporating categorical information (Kumar et al., 2025; Zhang et al., 2025) , even when such models do not inherently output calibrated uncertainty metrics. These stochastic data generation approaches could thus facilitate hotspot identification based on both impact significance and data confidence when combined with established uncertainty propagation approaches (e.g., Monte Carlo simulation) (Olivetti et al., 2013; Tecchio et al., 2019), allowing a highly uncertain parent

process to be disaggregated into a set of child processes that are smaller in impact magnitude, due to mass conservation, and potentially with higher data confidence, thereby providing a new pathway for system boundary expansion and refinement.

Here, we propose a conceptual, generative algorithm to quantitatively identify the system boundary in LCA by exploiting the tree analogy of product systems and leveraging uncertainty-aware inventory data, even when fully probabilistic inventories are not readily available. The algorithm begins by identifying the foreground system and then expands the system stepwise, guided by hotspot identification that accounts for both a process's impact significance and data confidence. To study the behavior of alternative boundary rules under controlled conditions, we introduce a numerical boundary-analysis framework that generates synthetic product-system trees with explicitly modeled uncertainty. Using this testbed, we systematically assessed how cutoff thresholds, data quality, and foreground impact concentration (i.e., the tendency of one or a few foreground components to dominate total impacts) affect the final system boundary complexity, the iteration steps required to establish system boundaries, and the truncation error. By coupling an explicit uncertainty-informed expansion rule with a reproducible stochastic benchmark, our proposed framework could potentially help advance the methodological transition from the current predefined, supply chain-based approach toward a more transparent, quantitative, and uncertainty-consistent methodology.

2. Materials and Methods

This section is structured into two main components. Section 2.1 details the mathematical and conceptual framework for the proposed iterative boundary expansion algorithm, including 1) the modeling setup step (i.e., foreground system definition and set initialization) (Figure 1-green), 2) the hotspot identification step (i.e., system definition, inventory analysis for the active set, Monte Carlo simulation, contribution calculation, and hotspot determination) (Figure 1-purple), and 3) system update or end of analysis (Figure 1-blue). Section 2.2 describes the simulation design and synthetic product system tree configurations used to evaluate the proposed algorithm's behavior under controlled conditions of data quality, foreground concentration, and cut-off thresholds.

Definitions of indices, sets, and parameters are listed in Table 1, with an example case study in the Supporting Information.

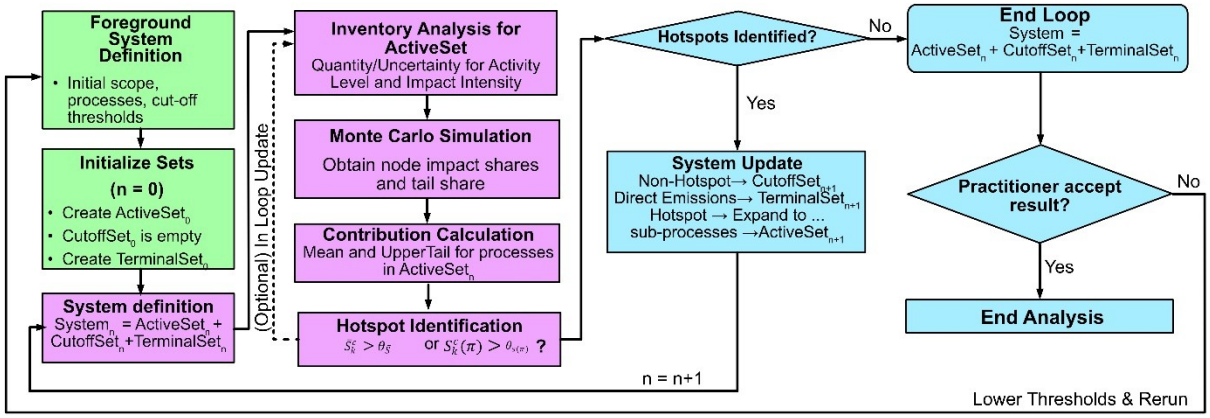


Figure 1. The workflow for the system expansion algorithm. Green boxes refer to modeling setup steps, purple boxes refer to hotspot identification steps, and blue boxes refer to system update or end of analysis.

Table 1. Definitions of indices, sets, and parameters for the method and examples using the homemade pizza illustrative case study.

Symbol	Definition	Example: Homemade pizza*
n	Progression of boundary refinement index, starting at $n = 0$ for the foreground system	$n = 0$ is the initial "Home Preparation" stage that includes processes from the questionnaire results
u	A process node	k = 14 for process "Households robot energy use (kWh)" and activity level A_{14} has a mean value of 0.111 and a standard deviation of 0.060
k	Index for a specific process	
A_k	Activity level for element k	
c	Index for a specific impact category	Global Warming Potential (GWP100)
EF_k^c	Emission factor of flow k for impact category c	kg CO_2e per kWh of electricity has a mean value of 0.154 and a standard deviation of 0.082
M	Number of Monte Carlo draws per run	10,000
$A\text{-}Set$	Set of currently active expandable processes at iteration n	Processes identified as hotspots at iteration n-1or foreground process for $n = 0$
$C\text{-}Set$	Set of cut off processes that were evaluated as no hotspots at iteration n	Processes that are kept aggregated because they are not important under the chosen hotspot rule
$T\text{-}Set$	Set of terminate processes that are not expandable (credit or direct emissions)	Final contributions that do not need further disaggregation
Sys_n	Modeled system representation at iteration n: $Sys_n = ASet_n \cup CSet_n \cup TSet_n$	The current modeled system boundary for the pizza calculation at iteration *0*.
$i_{k,n}^{(m,c)}$	Impact of process k in Monte Carlo draw m for category c	kg CO_2e emissions from process "Mozzarella cheese (kg)" in the sample m= 100 is 0.472
$I_{n,\lvert i^{(m,c)}\rvert i}$	Absolute total system impact at iteration n, Monte Carlo draw m for category c	Sum of absolute impacts at iteration m = 100 is 1.590
$s_{k,n}^{(m,c)}$	Share of contribution from process k at iteration n, Monte Carlo draw m, category c	At m = 100, process "Pizza baking energy use (kWh)" has a share of contribution of 0.154
$\theta_{\bar{s}}$	Threshold for hotspot identification based on mean share of contribution	Process "Mozzarella cheese (kg)" is identified as a hotspot process, as its $\overline{S}_k^{\square}$ is 0.261 and is larger than the predefined threshold $\theta_{\bar{s}} = i0.05$
$\overline{S}_k^c$	Draw-wise absolute shares (mean of $s_{k,n}^{(m,c)}$ over Monte Carlo draws M)	
π	Percentile used to compute $S_k^c(\pi)$	0.95
$\theta_{s(\pi)}$	Threshold for upper-tail-based hotspot identification	Process "Olive oil (kg)" is identified as a hotspot process as its $S_k^c(0.95)$ is 0.067 and is larger than the predefined threshold $\theta_{s(0.95)}$ = 0.05
$S_k^c(\pi)$	Contribution to quantile calculated as a quantile-based score at percentile π	

*A case study using a homemade pizza is shown in Supplementary Information.

2.1 The Proposed Framework

2.1.1 Goal, Scope, and Foreground System Definition

The model setup step in the proposed framework involves defining the foreground system by applying cutoff rules to elements such as capital goods or infrastructure and specifying hotspot analysis thresholds. In terms of scope, as in the current LCA workflow, the system boundary is defined by the research goal and the available time and effort budget. The key output of this stage is a well-defined set of flows, including technosphere inputs (i.e., material input, energy input, and transport activities) and elementary flows (e.g., direct emissions) located at the edge of the foreground system and forming the initial set of candidates for iteration in the system boundary expansion. (i.e., $ActiveSet_0$) (Figure 2a).

2.1.2 Step-by-step Iterative Hotspot Analysis

With a defined $ActiveSet_0$, the algorithm enters its main iterative loop. For a selected environmental impact category c, the workflow for each iteration n is as follows:

I. System Definition. At iteration n, the full system is defined as:

$$Sys_n = A\text{-}Set_n \cup C\text{-}Set_n \cup T\text{-}Set_n. \quad (1)$$

where $A\text{-}Set_n$ contains the active set that includes current expandable processes, which can be material/energy input, transportation, or elementary flows that are either defined in the foreground system (for n = 0) or expanded from previous hotspots (for n > 0); the $C\text{-}Set_n$ is a set of cut-off processes that were evaluated in a previous iteration but were not identified as hotspots at previous iterations; and $T\text{-}Set_n$ is a set of processes that are not expandable (i.e., of type of credit or direct emissions). Initialization uses $C\text{-}Set_0 = \emptyset$ (Figure 2a).

II. Inventory and impact analysis for the $A\text{-}Set_n$. This step involves the collection of activity level data (e.g., the amount and uncertainty of mass input, transportation distance, and energy usage) and the emission factors for the processes in the $A\text{-}Set$. For example, for each process u_k, the activity level A_k can be modeled as a normal distribution:

$$A_k \sim Normal\left(\mu_{A,k}, \sigma^2_{A,k}\right) \quad (2)$$

where $\sigma^2_{A,k}$ could represent measurement uncertainty. The process's emission factor for the impact category c can be modeled as a lognormal distribution for positive emissions, representing data uncertainty from inventory assessment:

$$EF^c_k \sim lognormal\left(\mu^c_{EF,k}, \left(\sigma^c_{EF,k}\right)^2\right) \quad (3)$$ as shown in Figure 2b. Any alternative distributional form (e.g., truncated normal, beta, gamma, empirical bootstrap distributions) can be substituted when more appropriate—the hotspot decision rule remains unchanged in structure.

III. Monte Carlo simulation. A Monte Carlo simulation with M runs is then performed on Sys_n. For each run, a value is sampled from the probability distribution of the activity level and the corresponding emission factor of every process, yielding process level impact $i^{m,c}_k$ that calculated as:

$$i^{(m,c)}_k = A^{(m)}_k \cdot EF^{m,c}_k \quad (4)$$

and absolute total system impact $I^c_{n,m}$ calculated as:

$$I_{n,|¿^{(m,c)}|} = \sum_{k \in Sys_n} |i^{(m,c)}_k| ¿ \quad (5)$$ as shown in Figure 2c.

IV. Contribution Calculation. Using the Monte Carlo outputs for each process for each sampling iteration m, the algorithm computes the share of contribution of process k in draw m as:

$$s^{(m,c)}_{k,n} = \frac{\left|i_k(m,c)\right|}{I_{n,|¿|(m,c)} \in [0,1], \sum_{k \in System_n} s^{m,c}_{k,n} = 1 ¿} \quad (6)$$

as shown in Figure 2d.

V. Hotspot Identification. Each process $u_k \in A\text{-}Set_n$ is evaluated against two predefined cutoff thresholds based on the contribution share $s^{(m,c)}_{k,n}$, i.e., the mean-based share of contribution $\overline{S}^c_k$ calculated as:

$$\bar{S}_k^c = E\left[s_{k,n}^{(m,c)}\right] \approx \frac{1}{M}\sum_{m=1}^{M} s_{k,n}(m,c) \quad (7)$$

and the upper-tail-based shared of contribution $S_k^c(\pi)$:

$$S_k^c(\pi) = P_\pi\left(s_k^{(1,c)}, \ldots, s_k^{(M,c)}\right) \quad (8)$$

where $P_\pi(\cdot)$ is the empirical π-percentile (e.g., π = 95 or 99). A process is identified as hotspot if

$$\bar{S}_k^c > \theta_{\bar{S}} \text{ OR } S_k^c(\pi) > \theta_{s(\pi)} \quad (9)$$

where $\theta_{\bar{S}}$ and $\theta_{s(\pi)}$ are two of the predefined threshold similar to the cutoff rule, as illustrated in Figure 2e and 2f.

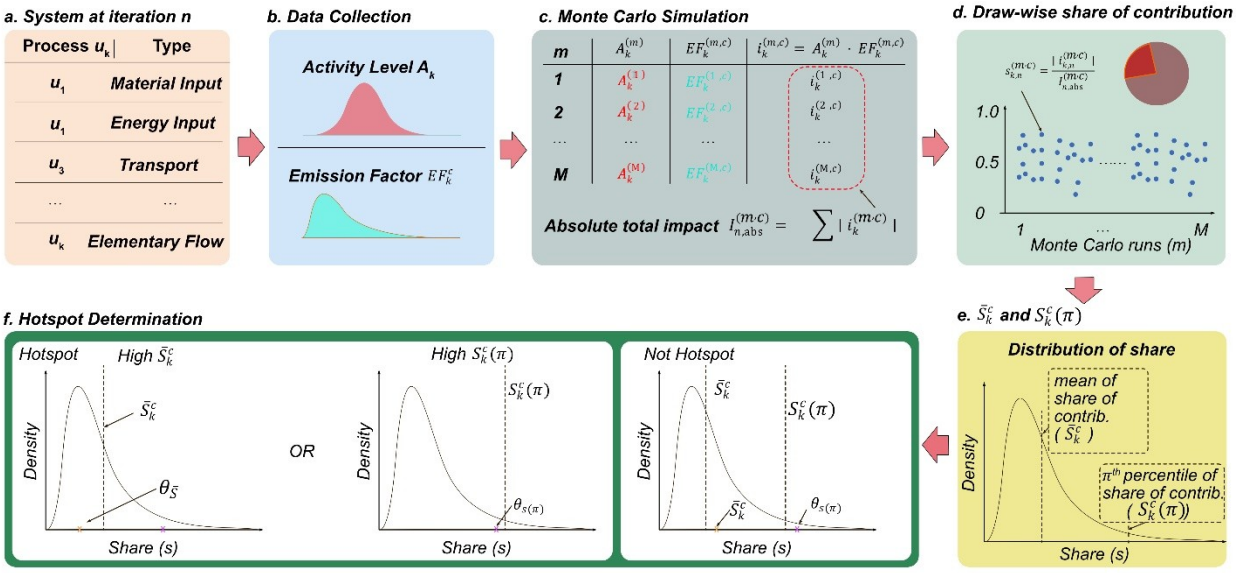


Figure 2. Workflow for an iteration step in the proposed system expansion algorithm.

Typically, in step II, the emission factors could be determined by using cradle-to-gate emission factors, including system processes in existing databases (e.g., ecoinvent (Wernet et al., 2016)), results of literature meta-analysis (e.g., the carbon footprint of food in different life cycle stages of foods reported in (Poore and Nemecek, 2018), and LCA-ready results (emission factors of transportation by different vehicles in the GREET model (Wang, 1999)), computed by proxying approaches, or using categorical information. Categorical variables (e.g., technology type, region, material/process class) can also be used to parameterize conditional distributions for uncertain inputs (e.g., conditional means and CVs for emission factors) (Dai et al., 2022). These conditional distributions can then be sampled within the Monte Carlo framework used by the hotspot criterion.

In step IV, the upper-tail-based share of contribute condition represents a high-end (upper-tail) estimate of process k's absolute impact-share under uncertainty. For example, for the threshold $\theta_{s(95)}$= 0.05, the identified hotspot processes' 95th percentile of share of contribution needs to be at least 5%. Thus, using upper-tail-based criteria identifies processes that may be important in the upper-tail of plausible outcomes driven by a large uncertainty level, even if their mean-based share of contribution (i.e., $\overline{S}_k^c$) is below its corresponding threshold. The thresholds are expected to be context-dependent (study goal, acceptable effort/expected accuracy), and the algorithm can be re-run with updated threshold settings to pursue higher resolution

2.1.3 Convergence Check and System Update

In the last step, we check whether any hotspot processes are identified under the selected criteria. If hotspots are present and the study goal/scope allows further resolution of upstream flows, the system can be updated: non-hotspot processes are moved from $A\text{-}Set_n$ to C-Set_{n+1}; hotspot processes are first disaggregated into their upstream constituent unit processes, and then put into $A\text{-}Set_{n+1}$, elementary flows are treated as non-expandable and moved to $T\text{-}Set_{n+1}$. Optionally, an in-loop update can be performed when early hotspot results are deemed inadequate for the intended analysis resolution (Figure 1). The algorithm is considered converged when no hotspots are identified. If additional resolution is desired, this can be pursued by lowering the hotspot thresholds and re-running the hotspot expansion within the same study boundary.

2.2 Simulation Design.

To demonstrate and examine the behavior of the proposed algorithm, we generated synthetic, rooted process trees with different configurations characterized by a set of parameters that control the data quality, hotspot concentration, and cutoff rules, as listed in Table 2. We first assign each foreground process an absolute mean impact from the sampled foreground shares and specify its uncertainty using pedigree-based coefficients of variation. We then expand hotspot nodes by generating child-process impacts whose draw-wise totals are constrained to the parent's empirical quantiles $\left[Q_{05}, Q_{95}\right]$and rescaled proportionally (i.e., children's total impact is calculated based on the parent mean and uncertainty), so upstream impacts inherit both the parent's magnitude

distribution and its uncertainty structure. This synthetic data was used to produce results shown in the subsequent section.

Table 2. Simulation configurations

Parameter		Value	Note
M		2,000	Number of uncertainty samples used to estimate process-level impacts and compute hotspot metrics
N_{run}		200	Number of independent synthetic tree realizations per configuration
N_0		10	Number of processes in the initial foreground system
l	Material (l_M)	[4,8]	Number of child processes
	Energy (l_E)	[2,4]	
	Transport (l_T)	[2,4]	
	Direct Emissions	1	
	Credit	Optional	
α		$\{0.55, 0.175$	Smaller values means more concentrated foreground contributions
b_t	Material	120	Sets the absolute impact scale for the initial foreground nodes; subsequent child impacts are generated by scaling parent impacts.
	Energy	80	
	Transport	60	
$\theta_{\bar{s}}=\theta_{s(\pi)}$		$\{0.01, 0.05, 0.15, 0.2\}$	Hotspot thresholds
π		$\{0.5, 0.6, 0.7$ $0.9, 0.95, 0.9$	Percentile level used to compute $S_k^c(\pi)$
P()		[Q_5, Q_{95}]	For each disaggregation, the draw-wise sum of child impacts is clamped to the parent's [$Q5$, $Q95$] interval (based on the parent's mean/uncertainty), then children are rescaled proportionally
CV	High	0.253	Pedigree-based coefficient of variation values (CV). Higher CV scenarios represents higher effort level in a LCA
	High-Median	0.281	
	Median	0.563	
	Median-Low	0.785	
	Low	1.532	

In detail, first, for the model initialization, the foreground system is designed with N_0 processes directly under a conceptual root node (i.e., the functional unit) (Figure 3a). The share of each node ω_k is sampled from a Dirichlet distribution:

$$(\omega_1, \ldots, \omega_{N_0}) \sim Dirichlet(\alpha, \ldots, \alpha) \quad (10)$$

So $\sum_{k=1}^{N_0} \omega_k = 1$. The concentration of the relative share among the N_0 initial foreground nodes is linked to the value of the archetype factor α: a smaller α generates more concentrated foreground contributions. The relative foreground shares of each node ω_k were then converted into absolute mean impacts of processes $i'_{k,0}$ by multiplying a type-specific base intensity $b_t(k)$ by the node's foreground share ω_k, with multiplicative lognormal noise ε_k to represent structural heterogeneity

$$\overline{I_{k,0}} = b_t(k) \cdot \omega_k \cdot \varepsilon_k, \varepsilon_k \sim Lognormal(0, \sigma_\varepsilon^2) \quad (11)$$

The constants $b_t(k)$ are selected for the synthetic testbed to impose a realistic ordering of typical magnitudes across process types and are not calibrated to any specific dataset. Then we sample the M-length Monte Carlo impact vector $i_{k,0}^{[m]}$ based on the uncertainty information of a process, which is assigned through a coefficient of variation (CV) following the widely used pedigree approach in the LCA community. Typically, the CV is sampled from a lognormal distribution centered on the pedigree mean CV for the selected data-quality scenario, with CV values obtained from Frischknecht et al.(2007). Energy and transport nodes are treated as one data-quality level better than the nominal scenario when applicable, to reflect better-characterized utilities and transport datasets typically. Each node's uncertain impact is then represented by an M-length Monte Carlo impact vector (Figure 3b and Figure 3c).

Second, for the expansion of each hotspot process, each expandable process can generate a fixed bundle of candidate children: l_M material-type children, l_T transport-type children, l_E energy child, one direct-emission child, and one optional credit child (Figure 3d). To ensure hierarchical

consistency (children collectively representing a plausible decomposition of the parent), as shown in Figure 3e and Figure 3f, we clamp the draw-wise sum of child impacts to the parent's empirical Monte Carlo quantiles $\left[Q_{05}, Q_{95}\right]$, and then rescale children proportionally so that the child total matches the clamped value. This preserves the relative composition among children while enforcing a bounded parent–child relationship.

Finally, we apply the iterative boundary-expansion algorithm and the stopping rules described in Section 2.1 and Figure 1 across a grid of cutoff thresholds θ and upper-tail percentiles π (see Table 2). For simplicity, we set the mean-share and tail-share cutoffs equal, i.e., $\theta_S = \theta_{s(\pi)} = \theta$. At each refinement iteration n, the algorithm updates the modeled system boundary by expanding nodes classified as hotspots and aggregating the remaining nodes into the cutoff set, while enforcing hierarchical consistency of parent–child impacts. This iterative procedure continues until the stopping rules in Section 2.1 are satisfied, and the final boundary is used to compute the hotspot metrics and truncation performance for each (θ, π) configuration.

To benchmark simplified boundary settings, we treat the most conservative configuration $(\theta = 0.01, \pi = 0.99)$ as a reference boundary. Performance of alternative settings is evaluated relative to this reference using common random numbers: each node ID is assigned a fixed stochastic stream so that differences across methods reflect boundary decisions (retention/expansion) rather than Monte Carlo noise. For each parameter setting we run N_{run} independent replicates (Table 2) and report outcome metrics including final-system node count, iteration depth, and truncation error.

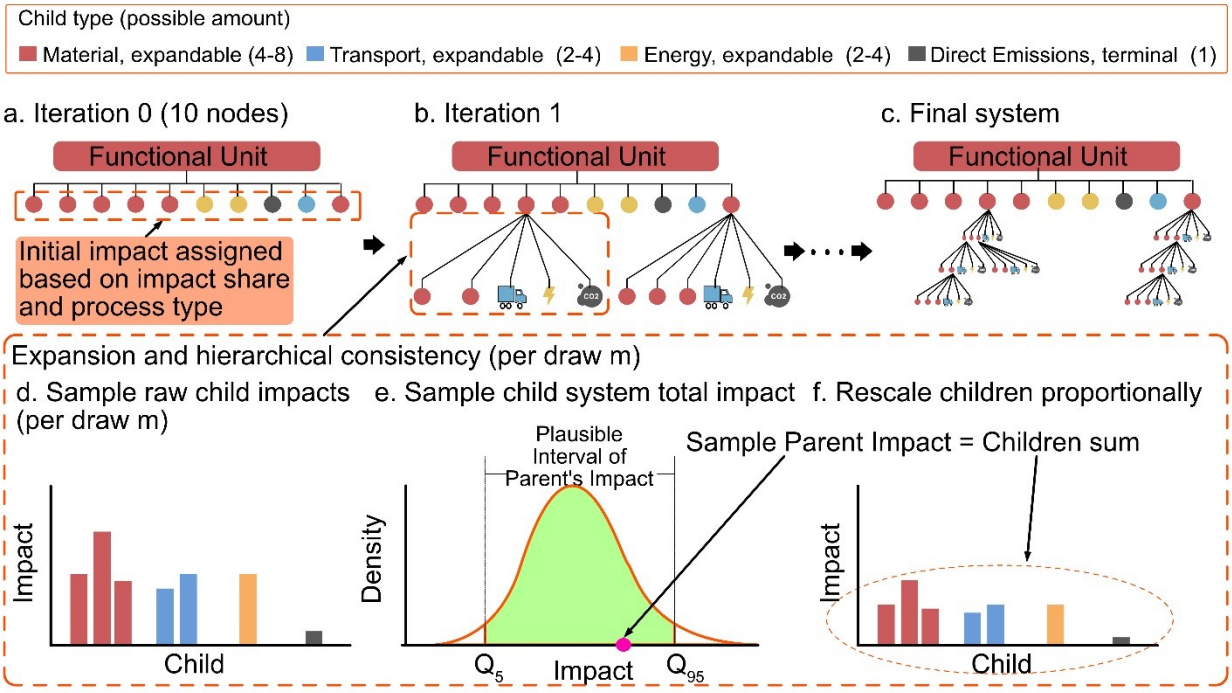


Figure 3. Illustration of the simulation of process expansion

3. Results

Overall, the results of the expansion behavior of the synthetic systems show that, under the selected synthetic configurations, 1) the applicability of the proposed algorithm for system boundary expansion and 2) the effectiveness of the stochastic simulation framework as a testbed for examining the impact of the system configurations.

First, the algorithm for the stochastic simulation of the life cycle system boundary expansion performs as designed. Our results show that the cutoff threshold θ (i.e., the predefined cutoff criteria for both mean share of contribution $\overline{S}_k^c$ and tail share of contribution $S_k^c(\pi)$) governs the speed and extent of system expansion: larger θyields fewer expansions, earlier stopping, and therefore smaller final systems; smaller θ sustains expansion over more iterations and transfers a larger fraction of total absolute impact from the active set into the cutoff and terminal sets. As shown in Figure 4a, the cumulative absolute-impact share carried by the active set decreases rapidly over the first iterations, while the cutoff and terminal sets first increase and then stabilize as the algorithm approaches termination until no remaining hotspots can be identified. This is consistent with the design that at each expansion step, nodes that do not meet the hotspot criterion are

transferred to the cutoff set, while nodes that meet the criterion are replaced by newly generated upstream children. Meanwhile, terminal contributions (e.g., direct emissions and credits) are immediately removed from the active set and accumulated after each expansion. Second, as shown in Figure 4b, the total number of nodes in the final system and the iteration count drop sharply as θ increases, as a result of fewer frontier nodes being identified as hotspots. Keeping θ fixed, the number of nodes that are identified as hotspots by the upper-tail-based rule as tail-percentile (i.e., π) increases, resulting in larger final boundaries and more iterations.

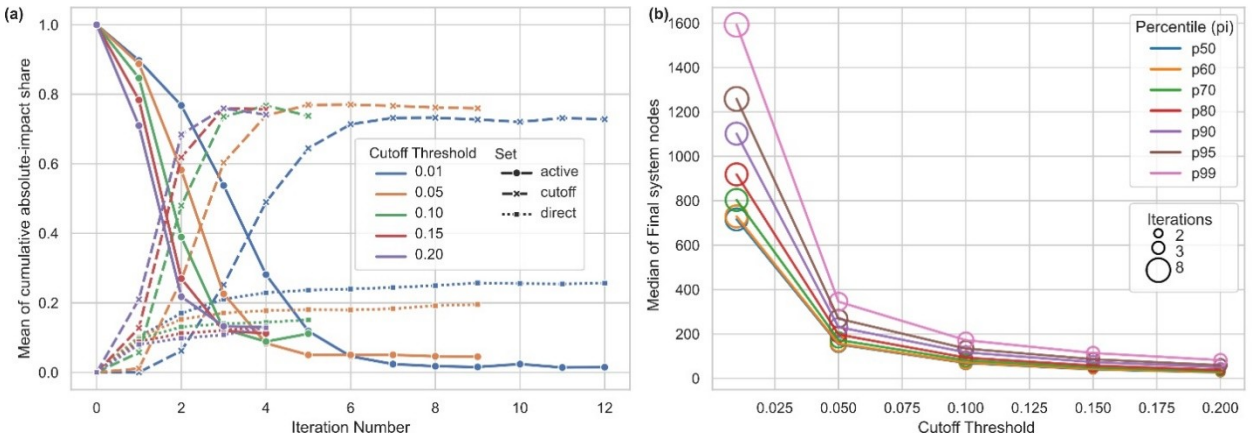


Figure 4. Convergence behavior under iterative thresholding. (a) Mean cumulative absolute-impact share by iteration for different cutoff thresholds (b) Median final system size (nodes) versus threshold θ across boundary percentiles π(marker size indicates iteration count).

Second, the overlay of the product system under different cutoff thresholds (θ) shows that where additional effort is incurred depends on the cutoff rule. Branch truncating can occur in the initial steps: For a concentrated system (archetype = 0.55, figure 5a) and a large cutoff threshold ($\theta = 0.20$, black dots), a few foreground system processes (i.e., the stars with no black dots connected) can be directly moved to the cutoff sets due to their small magnitude of impact. In addition, Foreground processes can also be truncated in a more evenly distributed foreground system when the cutoff threshold is large but via a different mechanism. As shown in Figure 5b, only 4 foreground processes were left for expansion in the foreground system with $\theta = 0.20$. This indicates that several of the excluded processes are comparable in magnitude but their individual mean contribution fall below the cutoff threshold, demonstrating that a coarse analysis with a high cutoff may not be appropriate for such a diffuse system.

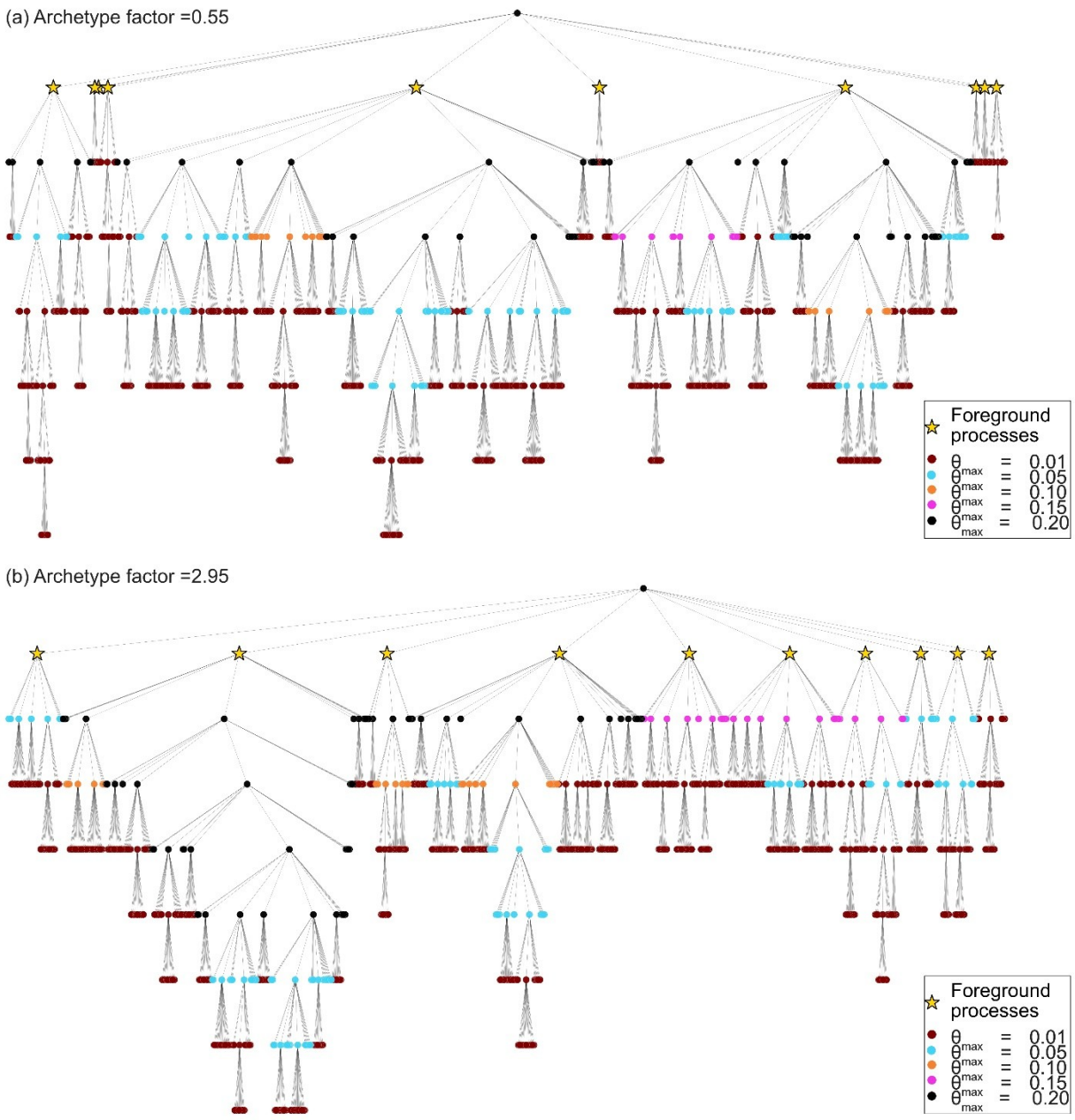


Figure 5. Layered system boundary graphs under two archetype settings (a concentrated foreground system (i.e., (a) foreground impact is dominated by a few key processes) with archetype = 0.55 and (b) a diffuse foreground system with archetype = 2.95). The baseline system with θ = 0.01 is plotted on the bottom layer with systems with a larger cutoff threshold stacked on top of it. A smaller threshold creates a more complete system that partially overlaps with the system generated based on a larger threshold, and the non-overlapping parts reveal where the additional effort is located. Data quality is median (CV = 0.563)

Lowering the threshold and applying a stricter cutoff rule, on the other hand, could create a large frontier and cause the total number of generated nodes to grow rapidly. This occurs because each retained parent process can generate multiple candidate child processes, which is consistent with the steep increase in total/final node counts in Figure 4b. The expansion of the system can follow a different pathway depending on the archetype of the foreground system: With more concentrated foreground contributions (lower archetype), system boundary refinement is pulled into a small number of dominant branches, producing deeper development along a few pathways (Figure 5a); With more diffuse foreground contributions (higher archetype), no single branch dominates and lowering the threshold distributes expansion across more regions of the network and admits additional secondary pathways, showing a broader and more distributed system boundary refinement (Figure 5b). Once the threshold is lowered, additional nodes appear outside the dominant flows (including deeper nodes with intermediate inclusion thresholds), indicating that applying the traditional main supply chain-only system boundary determination method could truncate cumulative background contributions, especially for diffuse systems where no single branch dominates.

At last, our results show that larger final systems are consistently associated with lower truncation error (compared to the reference system with high data quality, $\theta = 0.01$ and π=0.99), indicating a stable trade-off between the effort to expand the modeled boundary and the accuracy gained from increased coverage of total impact. As shown in Figure 6, lowering the cutoff threshold θ shifts results toward larger final-system node counts and systematically reduces truncation error. With a fixed θ, increasing the tail percentile π (from P50 to P99, colored circles) yields additional reductions in truncation error by adding the more potentially important but uncertain processes through the upper-tail-based rule. The cost of these gains depends strongly on data quality: under lower-quality data, between-run variability increases (shown as wider error bars) and achieving the same truncation error target requires substantially larger boundaries, implying a higher computational and data-collection burden. The truncation error can exceed the nominal net impact of the functional unit (100 units) because it represents the cumulative absolute mismatch across the entire system frontier. In systems with high data uncertainty and significant internal trade-offs (positive and negative impacts), the total absolute magnitude of unresolved flows often surpasses

the net system impact. In addition, since data quality can serve as an indicator of effort input level and the magnitude of truncation error as a measure of expected accuracy, the results suggest that selecting a harmonized expected accuracy, along with planned time and effort, and setting a cutoff threshold should be aligned. Limited data quality can significantly increase model complexity with only marginal benefits to model accuracy. It should be noted that the impact of the uncertainty-based rule could be enhanced by setting a stricter threshold for $\theta_{s(\pi)}$.

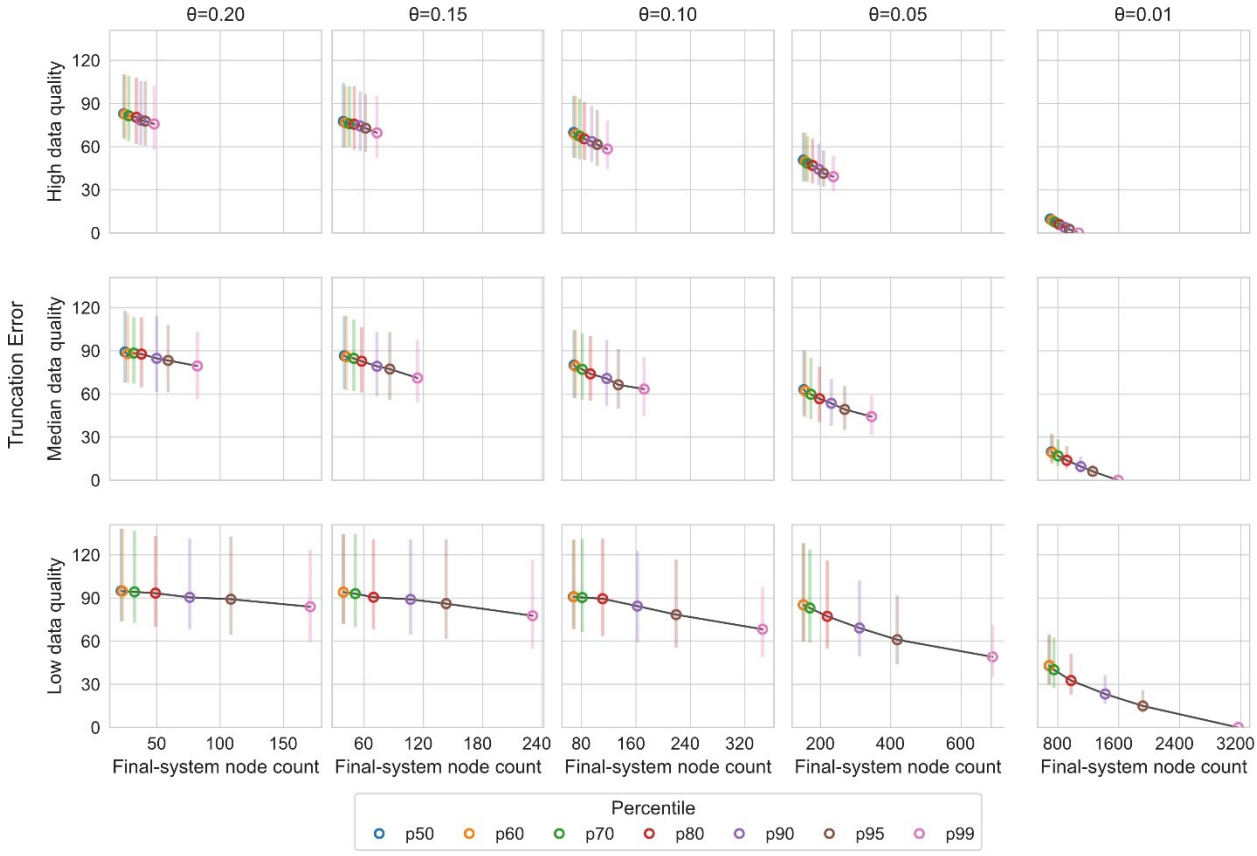


Figure 6. Effort–accuracy trade-off under boundary cutoff and tail rules. The y-axis measures the truncation error, calculated as the cumulative absolute mismatch of the simulated mean impact of the system compared to the reference impact ($\theta = 0.01$, $\pi = 0.99$) across all nodes in the system frontier. Because the root process is initialized with a baseline impact of 100 units, the numerical values on the y-axis represent the total magnitude of error or omitted impact relative to this initial system value.

4 Discussion

In this study, we proposed two complementary approaches, i.e., a system expansion algorithm built on the step-wise emission factor approach and uncertainty information, which iteratively decides whether each process should be expanded or kept aggregated based on uncertainty-informed hotspot identification (using both mean contribution and tail-share criteria), and a synthetic modeling framework that generates rooted process trees with explicitly modeled uncertainty, enabling controlled testing of the proposed algorithm across different cutoff thresholds, hotspot concentrations, and data-quality scenarios. Together, these two approaches provide both a formal method for quantitative system boundary expansion and a reproducible testbed to examine how boundary decisions affect convergence, model complexity, and truncation error under data limitations, and thus provide a solution to the question of which processes should be included and expanded, and which can be omitted. Depending on the study scope, the proposed algorithm can potentially be applied at three levels of expansion depth, including 1) expanding until convergence in a full, comprehensive LCA as designed, 2) terminating after iteration 1 and applying it to highlight the main uncertainty sources within the results for early-stage products or for eco-labeling using a simplified LCA method (e.g., food eco-labeling using ingredient emission factors (Dai et al., 2020b)), and 3) expanding only selected hotspots, similar to current main-supply-chain-based LCA practices, and using the proposed algorithm to justify or acknowledge the choices of system boundary.

The system expansion algorithm could potentially provide a way to document and report the intermediate iteration history of boundary selection decisions by explicitly connects the four phases of LCA—goal and scope definition, inventory assessment, impact assessment, and result interpretation—into a repeatable iterative workflow as each iteration recomputes a processes' contribution under uncertainty and uses uncertainty-informed hotspot results to determine which processes should be expanded to update the system boundary. LCA is designed as a dynamic, iterative method in ISO 14044, but such practices are not yet well-practiced or documented. For example, published LCA case studies commonly use a static, predefined system boundary, which could result in limited traceability and reproducibility across studies due to variability in how boundary decisions are made and reported (Bamber et al., 2019; Scrucca et al., 2020). Current

standards primarily require describing and documenting the final system boundary and any exclusions (Wolf et al., 2012). By making each iteration of boundary-expansion explicit and reportable, the structured approach could potentially help improve the transparency, reproducibility, and accessibility of a LCA study.

In addition, using the proposed approach could potentially help emphasize the integration of uncertainty analysis into LCA, transforming it from an optional—and often neglected—step into a core driver of the modeling process. Including the upper-tail-based cutoff rule first creates a placeholder for processes that practitioners have low confidence in, but that may still be important in the LCA analysis. Keeping all processes in the system could help address the paradoxical situation in the current impact-based cutoff method, where the negligibility of cutoff processes cannot be guaranteed without first measuring them. The algorithm can be applied in scenarios with both high acceptable uncertainty (e.g., initial screening and future-oriented LCA for products at the research and development stage) and low acceptable uncertainty (e.g., validating system boundary choice in a comprehensive LCA), depending on the defined goal and scope of the study. Typically, in early-stage eco-design applications, the framework's boundary prioritization could be combined with parametric LCA models by ensuring that parameter-driven optimizations are focused on the most relevant variables within a rigorously defined, uncertainty-adjusted system. This is aligned with recommendations to better characterize uncertainty and data granularity (Ross and Cheah, 2018) meanwhile still overcome the drawback of using mass or energy based cutoff rules, as in Li et al., 2014, as a low amount of input does not necessarily imply low impact.

The proposed synthetic testbed, on the other hand, provides a pathway to examine how key parameters, including the cutoff threshold, incorporation of uncertainty, and foreground impact concentration, affect system boundary expansion under controlled conditions where empirical studies are often limited by data and resource availability and modeling constraints. We quantitatively showed 1) that the cutoff threshold—the importance of which has long been acknowledged in LCA literature and standards (ISO, 2006a; Suh et al., 2004; Suh and Huppes, 2005; Tillman et al., 1994) —is the primary determinant of convergence and model complexity, 2) that tail-sensitive boundary expansion help reduce truncation error by retaining uncertain contributors, particularly under lower data quality, which is consistent with uncertainty-focused

perspectives on granularity effects on LCA outcomes (Ross and Cheah, 2018) and aligns with prior arguments that uncertainty-informed boundary selection can reduce the risk of invalid results from excluding upstream processes assumed to be negligible. For example, Suh et al., (2004) emphasize that the negligibility of cutoff processes cannot be guaranteed, and that small mass/energy does not necessarily imply negligible impacts. Importantly, although the final system size is expected to increase as the cutoff rule becomes more stringent (i.e., the threshold value decreases), our framework adds methodological value by formalizing the size–cutoff relationship and by incorporating an uncertainty-informed upper-tail hotspot criterion into the complexity control alongside the current solely mean-based criteria. Together, these elements help explain how tail-sensitive criteria can reduce truncation error, especially in low-data-quality scenarios, rather than merely producing different model sizes. Furthermore, we also show that the archetype mainly changes the dispersion of additional effort across branches, which suggests that boundary choices focused on the main supply chain may under-cover distributed background contributions, particularly in diffuse systems, enhancing understanding from hybrid LCAs showing that limited process-tree boundaries and truncations can miss meaningful contributions arising from broad upstream networks.

In addition, the synthetic simulation method enables detailed comparisons across parameter settings. As a result, differences in final boundary size, iteration count, and truncation error can be attributed to modeling choices (e.g., the cutoff thresholds and decision rules) rather than to uncontrolled differences. The platform generates synthetic but structurally plausible product-system trees with explicitly modeled uncertainty, executes alternative boundary rules, and records step-wise system composition (i.e., contributions from the active, cutoff, and terminal sets), expansion events, and outcome metrics. As such, the platform functions as a theoretical testbed for studying boundary-expansion behavior and benchmarking boundary-definition strategies, complementing broader methodological literature that often relies on case-based evaluation rather than fully controlled synthetic benchmarks (e.g., Bala et al., (2010); Douziech et al., (2021); Wang et al., (2021)). It should be noted that the quantitative findings using the synthetic simulation should be interpreted as characterizing how the proposed boundary-expansion logic behaves under the selected synthetic configurations (including the assumed tree-generation and uncertainty

structures), rather than as general properties of real LCA product systems. Therefore, generalization to applied LCA settings requires further empirical testing using real inventories and disaggregation resolutions consistent with the algorithm's expansion requirements.

Several limitations of the proposed framework need to be addressed in future research. First, the requirement for stage-wise emission factors, uncertainty information, and disaggregated child-node data complicates the integration of the proposed approach into the existing LCA framework. The algorithm relies on emission factors that represent a process's cradle-to-gate impact, which are only available for some sectors, e.g., the food system has open-source data for each of the life cycle stages as summarized in Poore and Nemecek (2018). In addition, uncertainty data is scarce in the current LCA literature. This data limitation makes our algorithm future-oriented. Future studies can use the proposed framework as a foundation for LCA automation by addressing two key steps: (1) automating process decomposition, which converts system processes into unit processes during system expansion, and (2) automating proxy data usage for hotspot identification. The first may be achieved using Large Language Models (LLMs) trained on technical manuals, patent descriptions, research literature, and existing databases, while the second could be enhanced using machine learning approaches. For example, LLMs or techniques such as Gaussian processes could systematically generate suitable proxies from product characteristics and quantify their associated uncertainty (Dai et al., 2022; Kumar et al., 2025; Zhang et al., 2025). Pursuing these advancements would streamline the LCA workflow and fundamentally enhance the replicability and transparency of the results. Notably, the present study is not based on a specific machine-learning/categorical pipeline; it uses synthetic uncertainty models to test the boundary-expansion logic.

Furthermore, a recognized limitation of the current framework is its primary focus on system boundary expansion; complex accounting scenarios—such as multifunctionality, multi-product allocation, and end-of-life recycling—were not addressed and remain areas for future research. Handling multifunctional processes (e.g., a refinery producing gasoline, diesel, and heavy oils) remains one of the unresolved methodological challenges in LCA (Guinée et al., 2021). The current main method is the allocation-based approach, which splits the multifunctional process into virtual monofunctional processes using a set of allocation factors. However, determining these factors often relies on subjective, value-laden choices, and choosing to partition a process based on

physical properties (e.g., mass or energy content) versus economic value can drastically alter the resulting environmental footprint of the coproducts (Guinée et al., 2021; Heijungs et al., 2021). Adapting the iterative, uncertainty-informed algorithm proposed in this study may eventually offer a natural computational alternative to subjective allocation. Specifically, the algorithm could systematically trace substituted market processes (e.g., alternative energy production) until their uncertainty-adjusted impacts fall below the cutoff threshold. This would allow a rigorous mapping of the avoided burden, completely bypassing the need for subjective partitioning factors, and would be consistent with international ISO standards for system expansion. Extending the framework to systematically incorporate and resolve these methodological challenges is an essential next step to bridge the gap between exploratory analysis and objective policy applications.

At last, it should be noted that significant negative emissions or highly permissive thresholds pose mathematical challenges to algorithm convergence. In typical LCAs, splitting a hotspot into child processes distributes the impact, leading to a natural stabilization where the algorithm stops finding new hotspots. However, convergence may become unstable if significant credits are present (e.g., direct air capture or bioenergy carbon removal and storage (Dai et al., 2025; Hunter-Sellars et al., 2026) because cancellation makes absolute contribution rankings non-monotonic. Very low thresholds may also cause the algorithm to chase borderline contributors deeper than practically useful. In scope-limited studies, convergence should be reframed as the stabilization of relevant hotspot changes, reported alongside Monte Carlo uncertainty within the truncated boundary.

Conclusion

This study introduces a conceptual, uncertainty-informed framework for iterative system boundary expansion in life cycle assessment (LCA) by providing a clear, repeatable workflow to determine which processes should be included and expanded, or omitted in a product system through quantitative rules of hotspot-based expansion and impact-based cutoffs, leveraging the inherently tree-shaped structure of upstream flows. Our methodology inherently encourages practitioners to document the evolution of the system boundary and to incorporate uncertainty analysis throughout the system boundary updating process, which could potentially improve the transparency, reproducibility, and accessibility of LCA studies. We recommend that practitioners

adopt this transparent approach by explicitly reporting iterative boundary decisions rather than predefining a main supply-chain-focused system boundary and only reporting the final exclusions. Future research should focus on streamlining this workflow by automating process decomposition and proxy data generation using advanced machine learning or large language models and by expanding the algorithm to handle complex accounting scenarios, including multifunctionality and computational solutions to allocation problems.

Supplementary Information

A case study demonstrating the system expansion algorithm, with full R code, is available in the Supplementary Information.

Acknowledgment

This material is based upon work at the Joint BioEnergy Institute supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Contract Number DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. This study was also supported by the U.S. Department of Energy, Critical Minerals and Energy Innovation, Bioenergy Technologies Office. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the U.S. government.

Declaration of Interests

C.D.S. has a financial interest in Cyklos Materials.

References

Bala, A., Raugei, M., Benveniste, G., Gazulla, C., Fullana-I-Palmer, P., 2010. Simplified tools for global warming potential evaluation: When “good enough” is best. Int. J. Life Cycle Assess. 15, 489–498. https://doi.org/10.1007/s11367-010-0153-x

Bamber, N., Turner, I., Arulnathan, V., Li, Y., Zargar Ershadi, S., Smart, A., Pelletier, N., 2019. Comparing sources and analysis of uncertainty in consequential and attributional life cycle assessment: review of current practice and recommendations. Int. J. Life Cycle Assess. 1–13. https://doi.org/10.1007/s11367-019-01663-1

Clark, M., Springmann, M., Rayner, M., Scarborough, P., Hill, J., Tilman, D., Macdiarmid, J.I., Fanzo, J., Bandy, L., Harrington, R.A., 2022. Estimating the environmental impacts of 57,000 food products. Proc. Natl. Acad. Sci. U. S. A. 119. https://doi.org/10.1073/PNAS.2120584119/-/DCSUPPLEMENTAL

Dai, T., Ellebracht, N.C., Hunter-Sellars, E., Aui, A., Goldstein, H.M., Li, W., Hellwinckel, C.M., Price, L., Wong, A.A., Nico, P., Basso, B., Robertson, G.P., Pett-Ridge, J., Langholtz, M., Baker, S.E., Pang, S.H., Scown, C.D., 2025. Land-based resources for engineered carbon dioxide removal in the United States exceed the expected needs. One Earth 8, 101349. https://doi.org/10.1016/j.oneear.2025.101349

Dai, T., Fleischer, A.S., Lee, R., Wemhoff, A.P., 2020a. Life cycle inventory regionalization and uncertainty characterization: A multilevel modeling approach. J. Clean. Prod. 242, 118459. https://doi.org/10.1016/j.jclepro.2019.118459

Dai, T., Jordaan, S.M., Wemhoff, A.P., 2022. Gaussian Process Regression as a Replicable, Streamlined Approach to Inventory and Uncertainty Analysis in Life Cycle Assessment. Environ. Sci. Technol. 56, 3821–3829. https://doi.org/10.1021/acs.est.1c04252

Dai, T., Yang, Y., Lee, R., Fleischer, A.S., Wemhoff, A.P., 2020b. Life cycle environmental impacts of food away from home and mitigation strategies—a review. J. Environ. Manage. 265. https://doi.org/10.1016/j.jenvman.2020.110471

Douziech, M., Ravier, G., Jolivet, R., Pérez-López, P., Blanc, I., 2021. How Far Can Life Cycle Assessment Be Simplified? A Protocol to Generate Simple and Accurate Models for the Assessment of Energy Systems and Its Application to Heat Production from Enhanced Geothermal Systems. Environ. Sci. & Technol. 55, 7571–7582. https://doi.org/10.1021/acs.est.0c06751

EC-JRC, 2012. Product Environmental Footprint (PEF) Guide. Eropean Commision Jt. Res. Cent. 154. https://doi.org/Ares(2012)873782 - 17/07/2012

Frischknecht, R., Jungbluth, N., Althaus, H., Doka, G., Dones, R., Heck, T., Hellweg, S., Hischier,

R., Nemecek, T., Rebitzer, G., Spielmann, M., Wernet, G., 2007. Overview and Methodology Data v2.0, ecoinvent report No. 1. Dübendorf.

Garcia, R., Freire, F., 2014. Carbon footprint of particleboard: A comparison between ISO/TS 14067, GHG Protocol, PAS 2050 and Climate Declaration. J. Clean. Prod. 66, 199–209. https://doi.org/10.1016/j.jclepro.2013.11.073

Guinée, J., Heijungs, R., Frischknecht, R., 2021. Multifunctionality in Life Cycle Inventory Analysis: Approaches and Solutions. pp. 73–95. https://doi.org/10.1007/978-3-030-62270-1_4

Heijungs, R., Allacker, K., Benetto, E., Brandão, M., Guinée, J., Schaubroeck, S., Schaubroeck, T., Zamagni, A., 2021. System Expansion and Substitution in LCA: A Lost Opportunity of ISO 14044 Amendment 2. Front. Sustain. 2. https://doi.org/10.3389/frsus.2021.692055

Hunter-Sellars, E., Dai, T., Ellebracht, N.C., Pilorgé, H., Pisciotta, M., Bump, A.P., Rodriguez Calzado, E., Hovorka, S.D., Scown, C.D., Pang, S.H., 2026. The Potential and Cost of Carbon Dioxide Removal Using Direct Air Capture with Land-Based Wind and Utility-Scale Photovoltaics. Environ. Sci. Technol. 60, 5430–5441. https://doi.org/10.1021/acs.est.5c14628

ISO, 2006a. ISO 14040:2006 Environment management-Life cycle assesment- Principles and framework.

ISO, 2006b. ISO 14044:2006-Environmental Management- Life cycle assessment- Requirements and guidelients.

Kumar, A., Nazemi, F., Kodamana, H., Ramteke, M., Bakshi, B.R., 2025. A Large Language Model-based Framework to Retrieve Life Cycle Inventory and Environmental Impact Data from Scientific Literature. Environ. Sci. Technol. 59, 22533–22543.

Li, T., Zhang, H., Liu, Z., Ke, Q., Alting, L., 2014. A system boundary identification method for life cycle assessment. Int. J. Life Cycle Assess. 19, 646–660. https://doi.org/10.1007/s11367-013-0654-5

Olivetti, E., Patanavanich, S., Kirchain, R., 2013. Exploring the viability of probabilistic under-specification to streamline life cycle assessment. Environ. Sci. Technol. 47, 5208–5216. https://doi.org/10.1021/es3042934

Poore, J., Nemecek, T., 2018. Reducing food's environmental impacts through producers and

consumers. Science (80-.). 360, 987–992. https://doi.org/10.1126/science.aaq0216

Rodrigues, C., Kirchain, R., Freire, F., Gregory, J., 2018. Streamlined environmental and cost life-cycle approach for building thermal retrofits: A case of residential buildings in South European climates. J. Clean. Prod. 172, 2625–2635. https://doi.org/10.1016/j.jclepro.2017.11.148

Ross, S.A., Cheah, L., 2018. Uncertainty Quantification in Life Cycle Assessments: Exploring Distribution Choice and Greater Data Granularity to Characterize Product Use. J. Ind. Ecol. 00, 1–12. https://doi.org/10.1111/jiec.12742

Scown, C.D., Horvath, A., McKone, T.E., 2011. Water Footprint of U.S. Transportation Fuels. Environ. Sci. Technol. 45, 2541–2553. https://doi.org/10.1021/es102633h

Scrucca, F., Baldassarri, C., Baldinelli, G., Bonamente, E., Rinaldi, S., Rotili, A., Barbanera, M., 2020. Uncertainty in LCA: An estimation of practitioner-related effects. J. Clean. Prod. 268, 122304. https://doi.org/10.1016/j.jclepro.2020.122304

Suh, S., Huppes, G., 2005. Methods for Life Cycle Inventory of a product 13. https://doi.org/10.1016/j.jclepro.2003.04.001

Suh, S., Lenzen, M., Treloar, G.J., Hondo, H., Horvath, A., Huppes, G., Jolliet, O., Klann, U., Krewitt, W., Moriguchi, Y., Munksgaard, J., Norris, G., 2004. System Boundary Selection in Life-Cycle Inventories Using Hybrid Approaches. Environ. Sci. Technol. 38, 657–664. https://doi.org/10.1021/es0263745

Tecchio, P., Gregory, J., Ghattas, R., Kirchain, R., 2019. Structured Under-Specification of Life Cycle Impact Assessment Data for Building Assemblies. J. Ind. Ecol. 23, 319–334. https://doi.org/10.1111/jiec.12746

Tillman, A.M., Ekvall, T., Baumann, H., Rydberg, T., 1994. Choice of system boundaries in life cycle assessment. J. Clean. Prod. 2, 21–29. https://doi.org/10.1016/0959-6526(94)90021-3

Wang, M., 1999. The Greenhouse Gases , Regulated Emissions , and Energy Use in Transportation (GREET) Model Version 1 . 5. Environ. Prot. 1–8.

Wang, Y., Levis, J.W., Barlaz, M.A., 2021. Development of Streamlined Life-Cycle Assessment for the Solid Waste Management System. Environ. Sci. Technol. 55, 5475–5484. https://doi.org/10.1021/acs.est.0c07461

Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., Weidema, B., 2016. The

ecoinvent database version 3 (part I): overview and methodology. Int. J. Life Cycle Assess. 21, 1218–1230. https://doi.org/10.1007/s11367-016-1087-8

Wolf, M.-A., Pant, R., Chomkhamsri, K., Sala, S., Pennington, D., 2012. The International Reference Life Cycle Data System (ILCD) Handbook, European Commission. JRC references reports. https://doi.org/10.2788/85727

Yue, D., Pandya, S., You, F., 2016. Integrating Hybrid Life Cycle Assessment with Multiobjective Optimization: A Modeling Framework. Environ. Sci. Technol. 50, 1501–1509. https://doi.org/10.1021/acs.est.5b04279

Zargar, S., Yao, Y., Tu, Q., 2022. A review of inventory modeling methods for missing data in life cycle assessment. J. Ind. Ecol. 26, 1676–1689. https://doi.org/10.1111/jiec.13305

Zhang, X., Guo, X., Zhao, J., Xiong, J., Tian, Y., 2025. Intelligent application of large language model to life cycle assessment methodology. J. Clean. Prod. 529, 146776.