Journal of Industrial Ecology
https://doi.org/10.1007/s44498-026-00170-5

RESEARCH

Robust policy evaluation in material flow analysis under deep uncertainty

Norbert Weijenberg[1,2] · Willem Auping[2] · Sietske Lensen[3] · Anna Schwarz[3,4] · Nils Thonemann[1]

Received: 29 January 2026 / Accepted: 14 August 2026

Abstract

This study proposes a novel framework for robust policy evaluation in material flow analysis (MFA) under deep uncertainty. It is motivated by our observation that, under data scarcity, conventional uncertainty analysis methods may overstate the available knowledge by attempting to characterize uncertainty beyond what limited empirical evidence can support. We propose a shift away from identifying the most probable policy outcomes and their associated uncertainty, toward identifying policies that remain reliable despite uncertainty. This is achieved by integrating MFA with exploratory modeling and analysis, which (i) characterizes broad, plausible uncertainty ranges to incorporate extreme scenarios and (ii) assesses policy performance across multiple robustness metrics to reflect diverse perspectives. In addition, the study presents methodological suggestions for handling the sum-to-one constraint of transfer coefficients, including a constrained elicitation procedure for defining uncertainty bounds, a sampling approach that preserves the sum-to-one property without requiring normalization, and a vulnerability analysis method that remains applicable despite this constraint. An illustrative case study demonstrates the operationalization of the framework and highlights its strengths and limitations under simplified conditions. Further research is needed to refine the methodology and enable a systematic application in practice.

Keywords Material flow analysis · Exploratory modeling · Deep uncertainty · Data scarcity · Decision-making framework · Policy

1 Introduction

Material flow analysis (MFA) emerged as an accounting tool to describe city metabolism and pollutant pathways (Hendriks et al., 2000). Nowadays, it is increasingly used to evaluate the potential impacts of material efficiency and circularity strategies (Lase et al., 2023; Pauliuk & Heeren, 2021; Slotte et al., 2025). This expanding application, coupled with the sparse, noisy, and diverse data typically underlying MFA, makes uncertainty analysis crucial for supporting risk-informed policy decisions (Liao et al., 2025; Maxim & van der Sluijs, 2011). Statistical approaches are regarded by Laner et al. (2014) as the most rigorous, but a large variety of uncertainty analysis methods have been applied to MFAs. Statistical approaches characterize uncertain input variables using functions such as probability distributions to represent likelihoods, or fuzzy membership functions to express degrees of truth (Laner et al., 2014). However, the shape of these distributions cannot reliably be inferred in situations where data points are scarce, which are precisely the situations in which uncertainty analysis is most critical.

Kawecki et al. (2018) presented a workaround to this problem in their analysis of European plastic flows, which has since gained traction in MFA studies on plastic pollution (Abbasi et al., 2023; Chen et al., 2023; Kawecki & Nowack, 2019). Their method defines probability distributions for transfer coefficients and material imports based on data quality indicators from the pedigree matrix of one or a few data points, followed by probabilistic uncertainty propagation (Kawecki et al., 2018).

✉ Nils Thonemann
n.thonemann@cml.leidenuniv.nl

[1] Institute of Environmental Sciences (CML), Leiden University, Leiden, The Netherlands

[2] Faculty of Technology, Policy, and Management (TPM), TU Delft, Delft, The Netherlands

[3] Netherlands Organization for Applied Scientific Research (TNO), Princetonlaan 8, 3584 CB Utrecht, The Netherlands

[4] LCA Department - Value Chain Team, CE Delft, Oude Delft 180, 2611 HH Delft, The Netherlands

Published online: 27 August 2026

However, the appropriateness of this probabilistic approach remains on several grounds exposed to critique. First, the probabilistic representation of uncertainty may overstate the available knowledge when data are insufficient in quantity to justify a likelihood-based interpretation (Clavreul et al., 2013). Second, the symmetric uncertainty bounds for single data points, resulting from a single data quality assessment, poorly reflect the typically skewed distributions of highly uncertain variables (Limpert et al., 2001). Third, data quality assessments are inherently subjective. Translating them into probabilistic uncertainty is often arbitrary, with key parameters left to the modeler's discretion (Laner et al., 2015a). Finally, this method may produce distribution tails that exceed physically meaningful bounds (i.e., 0–1 for transfer coefficients), necessitating truncation and disregarding a portion of the estimated uncertainty (Kawecki & Nowack, 2019).

Partial remedies exist to address these shortcomings. For instance, fuzzy membership functions can be used in place of probability distributions to avoid overstating the available knowledge (Laner et al., 2015b). Similarly, log-normal distributions can substitute for normal or linear ones to address asymmetric uncertainty (Laner et al., 2015a). However, in both cases, the reliability of the analysis remains limited by the quantitative arbitrariness of the data quality-based approach. Bayesian inference bypasses the need for data quality assessments by using prior distributions, based on expert judgment or default assumptions, which are iteratively updated as new data become available (Dong et al., 2023; Lupton & Allwood, 2018). Nevertheless, the appropriateness of probabilistic framing remains questionable when data is scarce.

This fundamental limitation, the inability to objectively and consistently define probability distributions in data-sparse settings, constitutes a form of 'deep uncertainty' (Lempert et al., 2003). Under such conditions, using a single MFA model with weakly supported probability estimates may misinform decision-makers. In response, Schwab and Rechberger (2017) propose a dimensionless uncertainty measure, enabling comparison of uncertainty levels across different MFAs or stages of model development. However, its abstract nature can hinder interpretation by stakeholders and provides limited insight in the absence of a reference model. In addition, it does not capture non-parametric sources of deep uncertainty, which arise when stakeholders disagree on model structures, system boundaries, or outcomes of interest (Lempert et al., 2003).

To address these limitations, this study proposes exploratory modeling and analysis (EMA), which systematically evaluates a wide range of model parameterizations and interpretive value systems (Agusdinata, 2008). Since EMA is a central component of robustness frameworks (Herman et al., 2015) and decision-making under deep uncertainty (Marchau et al., 2019), it is well-aligned with the increasing application of MFA to policy evaluation. EMA has already been applied in combination with simulation modeling methods such as agent-based modeling and system dynamics (Kwakkel & Pruyt, 2013, 2015). However, its integration with MFA remains unexplored.

In deeply uncertain MFA models, EMA offers three main advantages over probabilistic MFA for policy evaluation. First, by avoiding probabilistic framing, EMA explicitly acknowledges that probabilities may be unknowable (Herman et al., 2015). Second, by stress-testing policies across broad, plausible uncertainty bounds, EMA helps identify strategies that perform well under diverse and extreme conditions (Bankes et al., 2013; Lempert et al., 2003). Finally, EMA accommodates both parametric uncertainties, related to numerical inputs, and non-parametric uncertainties. The latter includes normative choices related to system definition (Laner et al., 2014), such as the inclusion or exclusion of specific material categories (e.g. rubber in a plastics-focused MFA). They also include network structure uncertainty, which concerns the inclusion or omission of processes and material flows within the MFA system (Liao et al., 2025). This addresses a key gap in MFA literature, where network structure uncertainty is seldom considered (Liao et al., 2025).

In this study, we developed a framework integrating MFA and EMA to evaluate policies under conditions of deep uncertainty. We tested a pilot implementation of the framework in a case study on plastic pollution from polyethylene terephthalate (PET) beverage bottles in the Netherlands. This case study was not intended to support decision-making, but rather to exemplify the proposed method and explore its strengths and limitations. Building on the MFA developed by Schwarz et al. (2023), we assessed the robustness of eight policy proposals addressing plastic pollution. Following Lempert et al. (2003), a policy was considered robust if it performed reasonably well relative to its alternatives across a wide range of plausible system behaviors and value perspectives. In addition, we identified the uncertain factors most responsible for a policy's failure to meet stakeholder requirements, also known as robustness controls (Herman et al., 2015).

2 Methodology

To develop a framework merging MFA and EMA, we pursued both a terminological and procedural integration of the two methodologies. Terminologically, we classified model components of MFA according to the XLRM framework, where XLRM denotes four categories of model components: external factors, levers, relationships, and performance metrics (Kwakkel, 2017; Lempert et al., 2003). This classification is illustrated in Fig. 1 and Table 1.

External factors encompass all uncertain numerical or categorical inputs, such as material imports, transfer coefficients, flow magnitudes, and product lifetimes. Conversely, constants represent the model variables that are not treated as uncertain, such as those fixed by the model structure or based on precisely known values (Kwakkel, 2017). For example, transfer coefficients inherently equal one when all incoming material is assigned to a single downstream process. Similarly, safety regulations may fix product lifetimes at one year through annual replacement rules.

Relationships capture non-parametric uncertainty in MFA and can be divided into scenario uncertainty, arising from normative choices, and model uncertainty, stemming from simplifications of the real-world system (Huijbregts et al., 2003; Laner et al., 2014; Lempert et al., 2003). Both types of relationships can be parametrized to enable sampling across alternative normative and modeling choices. Each unique configuration of external factors and relationships defines a scenario (Lempert et al., 2003).

Levers represent numerical or categorical variables influenced by policy (Lempert et al., 2003). However, unlike in traditional EMA, they cannot directly correspond to interventions in MFA. For example, a policy promoting the installation of reverse vending machines cannot directly change an MFA variable. Instead, its effect could be modeled by an increase in the share of collected bottles. Consequently, policy levers in MFA act on external factors, whereas in traditional EMA, levers and external factors are distinct (Kwakkel, 2017). A specific configuration of levers defines a policy alternative, while their absence represents the business-as-usual case. Together, a policy combined with a scenario constitutes an experiment.

Performance metrics are the indicators used to assess policy outcomes (Kwakkel, 2017), which in MFA can be classified into stock metrics (e.g., the accumulation of materials in sinks), flow metrics (e.g., the availability of secondary resources), and function metrics (e.g., recycling rates). These provide the foundation for robustness metrics, which summarize policy performance across scenarios using satisficing, regret, or statistical measures (Kwakkel et al., 2016).

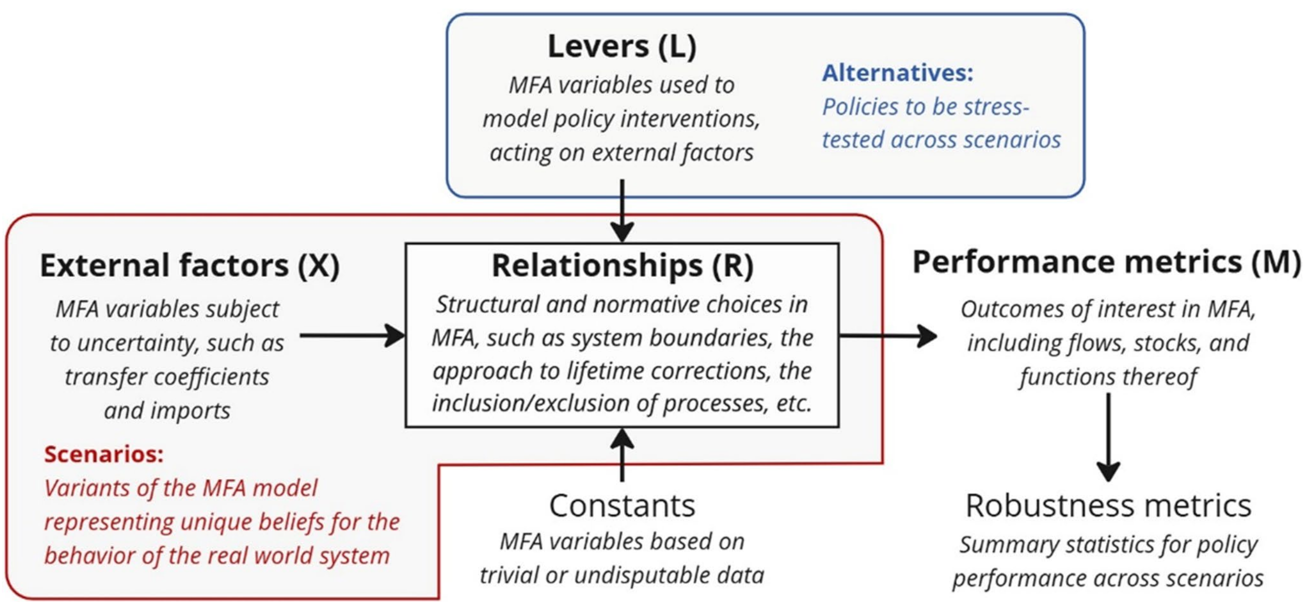

Fig. 1 Model structure of exploratory material flow modeling and analysis, based on the XLRM framework (Kwakkel, 2017)

Table 1 Examples of MFA model components, organized according to XLRM

External factors (X)	Levers (L)	Relationships (R)	Performance metrics (M)
■ Transfer coefficients ■ Material imports ■ Product lifetimes ■ Flow magnitudes	■ Changes a subset of X, depending on the policy	■ Time horizon ■ Inclusion/exclusion of certain processes, flows, and materials ■ Geographical and sectoral scope ■ Time step ■ Material grouping and aggregation	■ Accumulation of material in environmental sinks (*stock metric*) ■ Availability of secondary resource (*flow metric*) ■ End-of-life recycling rate (*function metric*)

Here, satisficing metrics count the number of scenarios meeting a minimum performance threshold, regret metrics compare policy performance to a reference performance in the same scenario, and statistical metrics describe the distribution of performance across scenarios using summary statistics (Kwakkel et al., 2016).

Building on this terminological integration, we procedurally combined the MFA research cycle (Brunner & Rechberger, 2016) with the taxonomy of robustness frameworks described by Herman et al. (2015), consisting of (1) the specification of policy alternatives, (2) the selection and sampling of external factors, (3) the quantification of robustness metrics, and (4) the identification of robustness controls. Several methodological decisions shaped this integration, based on the options outlined by Herman et al. (2015). First, we chose to prespecify policies rather than generate them by automated search, since policies in MFA rarely act directly on variables and must be translated into interventions with care. Second, we adopted global uncertainty sampling across wide, plausible ranges, consistent with exploratory modeling principles. Third, we left the choice of robustness metrics open to context-specific selection, rather than prescribing a universal measure, as recommended by Kwakkel et al. (2016). Fourth, we chose to identify robustness controls using a factor-ranking approach that evaluates the influence of uncertainties on policy performance (Herman et al., 2015). Three further design choices concerned the specification of uncertainty bounds, uncertainty propagation, and vulnerability analysis, discussed in more detail below.

For specifying plausible uncertainty bounds, we sought a method that establishes wide ranges that capture all credible values, thereby avoiding the risk of overlooking important vulnerabilities (Herman et al., 2015). In a data-scarce context, we consider it inappropriate to rely on extreme observed values or semi-quantitative methods. Instead, we adopted expert elicitation, a practice well established in MFA (Dong et al., 2023; Mathieux & Brissaud, 2010; Montangero & Belevi, 2007) and extended it with a novel procedure that respects the sum-to-one constraint of transfer coefficients, which ensures that the outflows of each MFA process sum to the total inflow. This ensured that uncertainty bounds on transfer coefficients were internally consistent without requiring post hoc adjustments, yielding a more faithful representation of expert judgment (Supplementary Information SI1 Section A).

For uncertainty propagation, uniform sampling was unsuitable because it conflicts with the sum-to-one constraint of transfer coefficients. While normalization could enforce this constraint, it risks distorting the underlying uncertainty distributions (Kawecki et al., 2018). Therefore, we adopted multivariate sampling from truncated, flat Dirichlet distributions (Ng et al., 2011), which preserve uniformity within the uncertainty bounds while accounting for the sum-to-once constraint. To approximate these distributions efficiently, we applied a two-blocked Gibbs sampling procedure (Park & Lee, 2022), as detailed in Supplementary Information SI1 Section B.

Finally, for vulnerability analysis (i.e., the process of identifying robustness controls), we adopted Spearman's rank correlation coefficient (Supplementary Information SI1 Section C). Spearman's rank correlation assesses the strength and direction of monotonic relationships between uncertain factors and policy performance. It is well suited to the constrained structure of the MFA model but may underestimate higher-order interactions (Cao et al., 2018; Wiedenhofer et al., 2019). Theoretically, variance-based methods could be adapted to accommodate the dependency structure of transfer coefficients and capture higher-order interactions. However, the interactions in MFA arise from multiplicative chaining of transfer coefficients, leading to network-wide interactions rather than low-order interaction terms. Capturing these effects would substantially increase methodological complexity while reducing interpretability of the analysis. Therefore, we use Spearman's rank correlation as a simpler, more interpretable approach.

3 Framework

We now present the framework integrating MFA and EMA (Fig. 2). Four of its stages are informed by the taxonomy of robustness frameworks (Herman et al., 2015), namely: the selection of decision alternatives (Stage 1), the selection of uncertain factors (Stage 3a), the quantification of robustness metrics (Stage 4a), and the identification of robustness controls (Stage 4b). The remaining stages are primarily derived from the MFA research cycle, covering the system definition (Stage 2), data collection (Stage 3b), and the illustration and interpretation of results (Stage 5) (Brunner & Rechberger, 2016).

3.1 Stage 1: Problem framing and policy definition

Stage 1 begins with formulating the problem statement in collaboration with policymakers and stakeholders. The problem must concern anthropogenic or environmental material flows and be characterized by deep uncertainty to ensure the relevance of the framework. Next, the analyst and policymakers define policy objectives, such as reducing plastic emissions, and translate these into MFA-type performance metrics, such as the cumulative flow of plastics into the ocean over a decade. Subsequently, they discuss appropriate performance thresholds to distinguish successful from failing policies and select robustness metrics to summarize policy performance across the full range of scenarios.

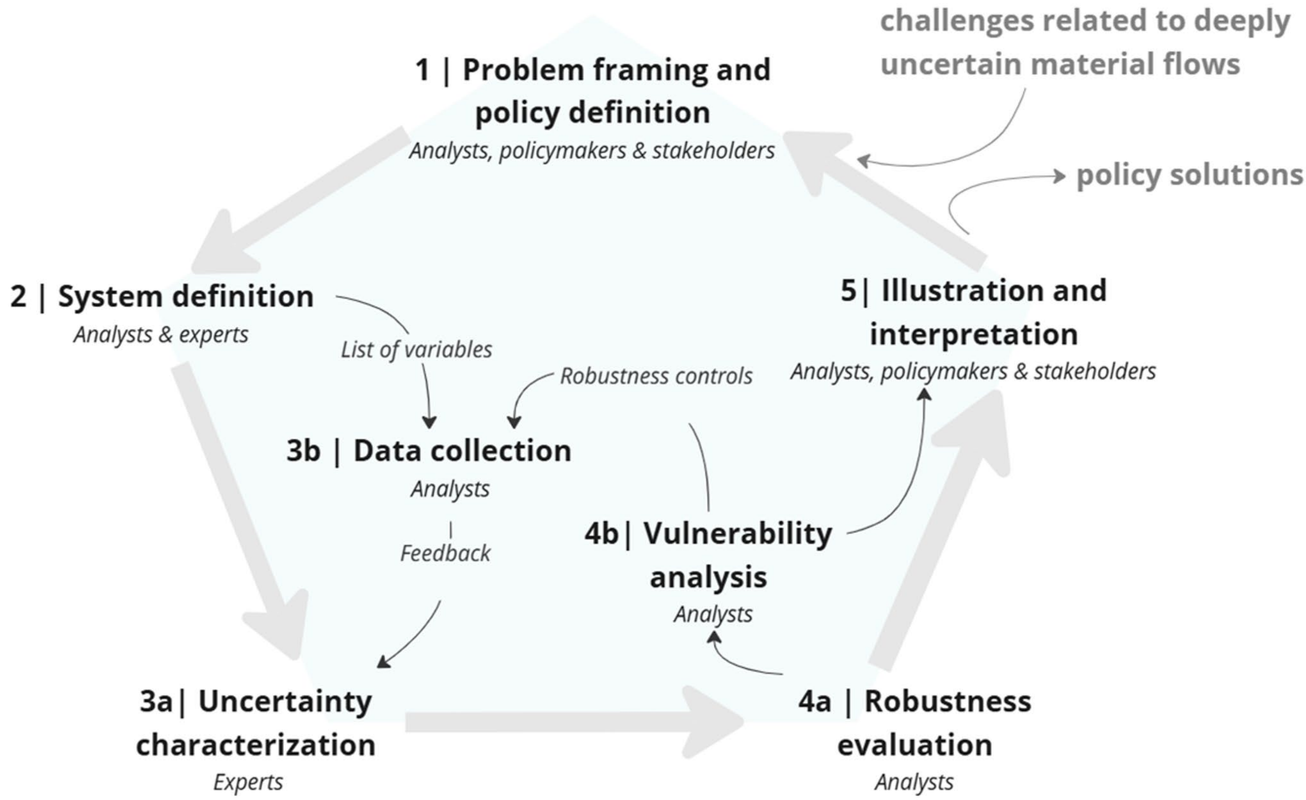

Fig. 2 Overview of the exploratory material flow modeling and analysis framework

Because no single robustness metric is universally superior, they should adopt context-relevant measures, apply multiple metrics in parallel, and synthesize their results (Kwakkel et al., 2016). This stage is concluded by specifying policy proposals for the analysis.

3.2 Stage 2: System definition

In Stage 2, the analyst defines the MFA system by specifying relevant materials, system boundaries, processes, and flows. The geographic and temporal scope typically follows from the problem statement, while the time step is selected to smooth out short-term fluctuations, often resulting in annual intervals for anthropogenic systems (Brunner & Rechberger, 2016). To identify key flows and processes, the analyst may consult academic literature, national reports, and subject-matter experts. Experts then review the qualitative MFA system, while the analyst documents any disagreements as uncertain relationships. In principle, all independent variables in the MFA model are treated as external factors, except those fixed by the model structure or known with certainty, which are treated as constants. Experts conclude this stage by estimating the expected effect of each policy on the MFA system, enabling the quantitative integration of policy levers into the MFA model. These estimates may be based on expert judgment, empirical evidence, or more detailed modeling approaches, such as system dynamics modeling, depending on data availability and the intended depth of analysis.

3.3 Stage 3: Uncertainty characterization and data collection

In this stage, the analyst designs a survey to elicit plausible uncertainty bounds for each external factor from subject-matter experts. The MFA system may be divided into thematic subsystems, enabling elicitation with more specialized experts and minimizing the time burden on each participant. Preparing the survey involves translating external factors into practical questions, such as: What percentage of consumer PET bottles used in the Netherlands in 2024 was used on the go? Experts are then asked to report the lowest and highest values they find able to be believed (Stage 3a). To elicit transfer coefficients under the sum-to-one constraint, the analyst may use the Excel spreadsheet available on GitHub (https://github.com/nweijenb2001/emfma/tree/main/elicitation), which applies the constrained elicitation procedure described in Supplementary Information SI1 Section A.

Before elicitation, the analyst should aim to gather high-quality data to support more informed expert judgments

(Stage 3b). However, this information should be withheld until a second elicitation round to minimize anchoring bias (Furnham & Boo, 2011). Data collection should focus especially on robustness controls if vulnerability analysis has been conducted previously.

When multiple experts provide uncertainty bounds for the same transfer coefficient, the analyst could aggregate them by taking their union. This would align with EMA's preference for broader uncertainty bounds, which may include extreme scenarios where most policies fail, but lower the risk of overlooking critical vulnerabilities (Herman et al., 2015). However, this approach would demand post-aggregation adjustments, such as trimming (Supplementary Information SI1 Section D), to ensure compatible ranges for transfer coefficients. These adjustments are undesirable because they weaken the direct translation of expert judgment into uncertainty bounds. We therefore recommend employing group elicitation approaches such as the SHELF protocol (Soares et al., 2024), which enable consensus estimates for uncertainty bounds without aggregating individual expert inputs.

3.4 Stage 4: Robustness evaluation and vulnerability analysis

In Stage 4, the analyst generates a large ensemble of scenarios by sampling across uncertain relationships and within plausible uncertainty bounds for external factors. Transfer coefficients can be sampled using the two-blocked Gibbs sampling procedure detailed in Supplementary Information SI1 Section B to respect the sum-to-one constraint, while independent external factors, such as material imports, can be sampled uniformly or via pseudo-random methods. Once the scenarios are generated, the analyst evaluates the performance of each policy in each scenario by solving the MFA system. Then, in Stage 4a, the analyst quantifies policy robustness using the chosen robustness metrics. Finally, in Stage 4b, the analyst applies Spearman's rank correlation coefficient to identify the vulnerabilities of each policy concerning each performance metric.

The analyst typically determines the appropriate number of scenarios through trial and error or by monitoring when robustness and vulnerability scores stabilize as the sample size increases (Kwakkel, 2017). If computational constraints prevent such convergence analysis, a limited set of diagnostics can be used as practical alternatives, though they do not replace proper convergence testing. In particular, a t-distribution-based test can be used to filter vulnerability correlations likely due to chance (Virtanen et al., 2020), while agreement in robustness scores between a limited number of duplicate model runs can be quantified using the Intraclass Correlation Coefficient (Koo & Li, 2016).

3.5 Stage 5: Illustration and interpretation

In the final stage, the analyst collaborates with policymakers and stakeholders to visualize and interpret the results. This process may uncover previously unrecognized challenges or prompt policy refinements, potentially initiating a new research cycle. Policy performance across the full range of scenarios can be illustrated using strip plots, while parallel coordinate plots can depict policy robustness across multiple metrics. Additionally, Sankey diagrams can visualize the MFA model under best- and/or worst-case conditions for each policy-performance metric pair. In these diagrams, flow colors can represent the Spearman's rank correlation coefficients of transfer coefficients or material imports, highlighting vulnerabilities unique to each policy and performance metric. Section 4.5 presents examples of these visualizations from the illustrative application.

4 Illustrative application of the framework

We applied a pilot version of the proposed framework to the plastic littering MFA developed by Schwarz et al. (2023) We refer to this as a pilot application because the Schwarz et al. model is extensive, not fully modularized, and actively under development, which limited our ability to demonstrate all aspects of the framework, in addition to constraints in expert availability. Specifically, group elicitation was replaced by elicitation with a single expert, policy effects were not elicited from experts, and uncertainty analysis was restricted to transfer coefficients.

Given these limitations, the case study is not intended to inform policy decisions, but instead to highlight several strengths and limitations of the framework, illustrate how its results can be visualized and interpreted, and demonstrate how it differs from conventional uncertainty analyses. The Python script utilized in the execution of the case study is available on GitHub (https://github.com/nweijenb2001/emfma/blob/main/main.py).

4.1 Problem framing and policy definition

In Stage 1, we selected the littering of PET consumer bottles in the Netherlands as the central problem. While the Dutch government has taken steps to reduce beverage bottle litter, the associated collection targets have not yet been achieved, highlighting the continued relevance of, this issue (Schep et al., 2025). Although policymakers were not directly involved in the definition of policy objectives, we designed them to align with the national targets for 2030. The selected policy objectives were (1) reducing microplastic emissions (Rijksoverheid, 2023) and (2) increasing PET recycling (Rijksoverheid, 2021). Corresponding flow-type performance metrics were defined for the period from 2025 to 2030 in the Netherlands as: (1) cumulative environmental microplastic emissions caused by consumer PET bottle consumption, and (2) cumulative secondary material production from used consumer PET bottles (i.e. macroplastic recovery flows). The first metric represents a minimization problem, where lower scores indicate better policy performance, while the second represents a maximization problem.

To evaluate policy robustness, we selected five robustness metrics based on Kwakkel et al. (2016), including one satisficing, one regret, and three statistical metrics. Each was designed with two variants to match the direction of the performance metric (minimization or maximization) and enhance clarity when plotting multiple metrics in a single figure. One of the metrics, the domain criterion, required threshold values to distinguish successful policies from failing ones. For environmental microplastic emissions, we defined success as cumulative emissions remaining below 70% of the business-as-usual level. This reflects the national target for 2030 (Rijksoverheid, 2023), though applied here specifically to PET bottles. For secondary material production, we defined success as achieving a secondary material production greater than 30% of PET bottle consumption, aligning with the national recycled content target for 2030, assuming closed-loop recycling (Rijksoverheid, 2021). Supplementary Information SI1 Section E provides a more detailed overview of the metrics.

To identify viable policy options, we consulted the study by Quik et al. (2024), which included an expert workshop to estimate the feasibility and effectiveness of various mitigation measures targeting plastic pollution in the Netherlands. We included all measures relevant to macroplastic emissions, excluding those unrelated to the context of PET bottles (e.g., banning EPS plastics in packaging) and those that would produce equivalent effects in the MFA model. For example, both a broader rollout of deposit systems and increased return system incentives would be represented in the model as an increase in the flow of consumer bottles to the deposit refund scheme. As these measures target the same transfer coefficients within the model and thus produce equivalent insights, we included only one of them, selecting the option with the higher combined average score for feasibility and effectiveness. This resulted in the selection of eight policies for evaluation alongside the business-as-usual case, as detailed in Table 2.

4.2 System definition

In Stage 2, we defined the qualitative MFA system, selected external factors, and specified policy levers. The MFA system was based entirely on the Python-based model of Schwarz et al. (2023), tracing yearly plastic flows from consumption to grave, including releases to the environment and subsequent dispersion and degradation. The model used an annual time step and distinguished between macroplastics (> 5 mm) and microplastics (< 5 mm) (Peano et al., 2020). We defined the geographic and temporal scope as the Netherlands for the period 2025–2030 to align with the performance metrics, and we focused exclusively on PET pathways from consumer beverage bottles within the model. Since PET bottles are single-use with an average estimated lifetime of one year, no lifetime corrections were necessary, consistent with a static MFA approach (Laner & Rechberger, 2016). We used a throughput approach without exports, meaning that all material entering the system accumulated in the model's sinks.

Since the Schwarz et al. model is not fully modularized, incorporating model uncertainty in addition to parametric uncertainty would have required substantial modifications to the original code. In addition, the elicitation burden on experts would have been too high without simplifications, given the large number of MFA processes involved. Therefore, model relationships, material import, and a subset of transfer coefficients were treated as constants in this

Table 2 Overview of policies included in the illustrative application, including scoring on feasibility (1-Easy to 3-Difficult) and effectiveness (1-High effectiveness to 3-Low effectiveness) as rated by 7–10 experts (Quik et al., 2024)

Policy	Description	Feasibility (1–3)	Effectiveness (1–3)
BAU	Business-as-usual, continuing system behavior of 2024	–	–
Reduce	Further restrict single-use disposable plastic packaging and items	1.78	1.11
Capture	Capture MP at production and recycling plants	1.33	1.56
Clean	Litter clean-up (roadsides, parks, woodlands)	1.50	2.00
Return	Return systems, incentives/Increase deposit systems	1.89	1.67
Homogenize	Reduce material complexity to reach more true circularity	2.67	1.44
Manage	Proper waste management systems/Improved waste management	2.22	1.89
Redesign	Redesign polymers to minimize shedding of MP	2.71	2.00
Extract	Extraction of plastics from landfills	2.80	2.40

demonstration. These simplifications limit the completeness of the analysis and would not be appropriate in a case study intended to inform policy insights.

The material import to the MFA, representing PET consumption levels, was taken from the original model (Schwarz et al., 2023). Some transfer coefficients were fixed because they were fully determined by the model structure (i.e., processes with a single inflow and outflow). Where transfer coefficients fulfilled equivalent roles, they were modeled as separate external factors but elicited only once to avoid redundancy. For example, the distribution of wastewater across treatment processes was assumed to be identical for wastewater containing microplastics and macroplastics. Lastly, a small number of transfer coefficients were treated as constant due to their presumed limited influence given their position in the MFA model, which is a limitation of this case study (see Supplementary Information SI1 Section F).

After defining the qualitative system, we hypothesized which external factors would be affected by each policy listed in Table 2. In the absence of expert input or detailed policy specifications, we applied a uniform ± 10% adjustment to the affected factors. This assumption was not intended to represent policy effectiveness realistically, but to enable a simple assessment of how changes in transfer coefficients affect system behavior. For example, the Return policy was modeled as a 10% increase in the deposit rate of on the go and at-home PET bottles relative to the business-as-usual case. Note that this reflects a proportional change in the transfer coefficient, not an absolute increase of 10 percentage points. To maintain mass balance, we offset increases in one coefficient with proportional reductions across others within the same process. Furthermore, we constrained adjusted coefficients to lie between 0 and 1, ensuring mathematical consistency. A complete overview of the policy levers is provided in Supplementary Information SI1 Section G.

4.3 Uncertainty characterization and data collection

In Stage 3, we designed an Excel-based survey to elicit plausible uncertainty bounds for the external factors. Following the procedures in Supplementary Information SI1 Section A, the survey ensured expert responses adhered to the sum-to-one constraint of transfer coefficients from the outset. It comprised 117 questions grouped into 43 MFA processes, asking participants to specify the lowest and highest plausible values for transfer coefficients under the business-as-usual case. Due to the limited availability of plastic pollution experts during this research, the survey was conducted with a single methodologically knowledgeable expert. This was acceptable for demonstrating the approach but not sufficient for policy-oriented analysis, which should involve a broader and context-dependent range of domain experts. Although data on external factors were available from the original model (Schwarz et al., 2023) and a similar study focused on the Netherlands (Quik et al., 2024), we deliberately withheld this information from the expert due to their frequent reliance on single measurements, proxies, and assumptions. Still, approximately 74% of the elicited bounds encompassed the original values used by Schwarz et al. (2023). The survey response is provided in Supplementary Information SI1 Section H.

4.4 Robustness evaluation and vulnerability analysis

In Stage 4, we used two-blocked Gibbs sampling to generate near-uniform samples within the bounds of the 136 external factors. The number of external factors is smaller than the number of survey questions because some factors recur in different parts of the system and were therefore elicited only once, as described earlier. To address the computational demands posed by running the full Schwarz et al. (2023) model, we limited the sample size to 500 scenarios. We then applied each of the nine policy alternatives to the scenarios, producing 4500 experiments. We determined policy performance in each experiment and quantified policy robustness according to the preselected metrics. Subsequently, we computed Spearman's rank correlation coefficients to explore the relationship between external factors and policy performance, applying a t-distribution significance test ($\alpha = 0.05$) to exclude correlations likely due to chance. To ensure stability, we repeated the sampling and analysis twice and quantified agreement between runs using the intraclass correlation coefficient (ICC), following the guidelines of Koo and Li (2016). The results of this agreement test are included in Supplementary Information SI1 Section I.

4.5 Illustration and interpretation

In Stage 5, we visualized the results of the first model run in terms of policy performance, robustness, and vulnerability. As stressed earlier, these results should not be interpreted as guidance for plastic pollution policy, but rather as an illustration of how the proposed framework can be applied, visualized, and interpreted.

Figure 3 illustrates policy performance with respect to reducing environmental microplastic emissions (Fig. 3a) and increasing secondary material production (Fig. 3b). The dots represent different scenarios, and a few lines are shown to indicate that policies were evaluated across the same set of scenarios. Lower scores indicate better outcomes in Fig. 3a but worse outcomes in Fig. 3b. Both plots show a wide range of plausible outcomes: approximately 10–130 tonnes for microplastic emissions and 50–550 tonnes for secondary

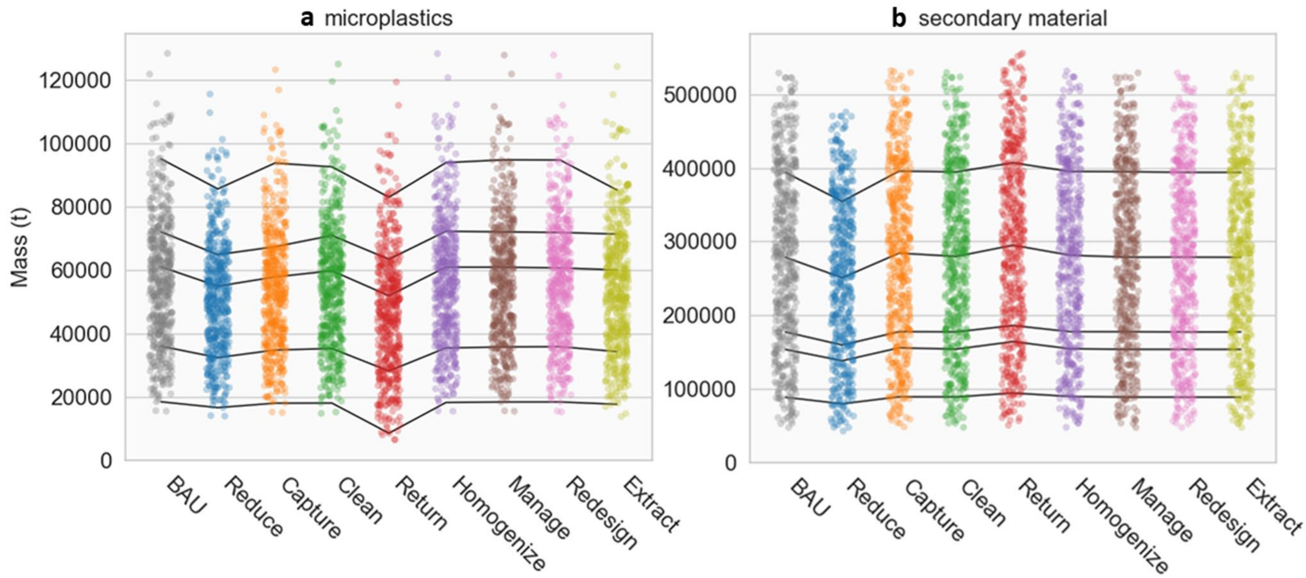

Fig. 3 Policy performance across scenarios in terms of **a** cumulative environmental microplastic emissions and **b** cumulative secondary material production, related to consumer PET bottles in the Netherlands from 2025 to 2030. The Python script used to generate this figure is provided on https://github.com/nweijenb2001/emfma/blob/main/main.py

material production. Distinct skew patterns are also evident: for example, microplastic emissions are more dispersed in the undesirable range than in the desirable range. The figure also enables preliminary policy comparisons. For instance, the *Return* policy, featuring deposit-refund incentives, achieves the best best-case outcome for microplastic emissions. Conversely, the *Reduce* policy produces the best worst-case outcome for microplastic emissions but the worst best-case outcome for secondary material production. This trade-off is expected, as reduced consumption limits the material available for recycling. The substantial overlap in policy outcomes reflects our use of broad uncertainty bounds, which aligns with the study's aim to assess robustness across diverse conditions.

Figure 4 presents policy robustness with respect to microplastic emissions (Fig. 4a) and secondary material production (Fig. 4b). Lower scores indicate greater robustness in Fig. 4a but lesser robustness in Fig. 4b. The ICC test indicated good to excellent agreement between robustness scores across the two model runs, but only after min–max normalization (Supplementary Information SI1 Section I). In other words, the relative robustness of policies remained consistent between two runs, whereas their absolute scores were unstable at a sample size of 500 scenarios.

The intersecting lines highlight trade-offs between policies related to the choice of robustness metric. For microplastic emissions, the *Return* policy achieved the lowest failure rate but exhibited a more severe worst-case outcome than *Reduce*, indicating that *Reduce* may be more risk-averse despite its higher failure rate. For secondary material production, *Return* was consistently the most robust across all metrics, while *Reduce* was the least robust. This pattern arises because lower PET consumption directly reduces the volume of end-of-life material available for recycling. The *Capture* policy, focusing on recovery at recycling plants, demonstrated moderate robustness for both policy objectives. In contrast, the *Extract* policy, targeting landfill recovery, achieved moderate robustness for microplastics emissions but ranked among the least robust for secondary material production. This is explained by the relatively small PET flow to landfills.

Figure 5a, b depict statistically significant vulnerabilities in the material flow system for environmental microplastic emissions and secondary material production, respectively. Insignificant correlations are shown in gray, appearing visually similar to non-correlations. The ICC test indicated good to excellent agreement between vulnerability scores across the two model runs. As vulnerabilities were largely independent of policies, only the business-as-usual case is presented, depicting the worst-case scenario for each performance metric in 2025. Environmental flows beyond human control, including transport processes such as water transport, burial or sedimentation, are omitted to improve the clarity of the figure. In the diagrams, red indicates outflows where transfer coefficients improve performance when large, while blue indicates outflows where larger values worsen outcomes.

The results indicate that the uncertainty in the share of deposited PET bottles strongly influences performance. Additionally, the proportions of plastic recycled,

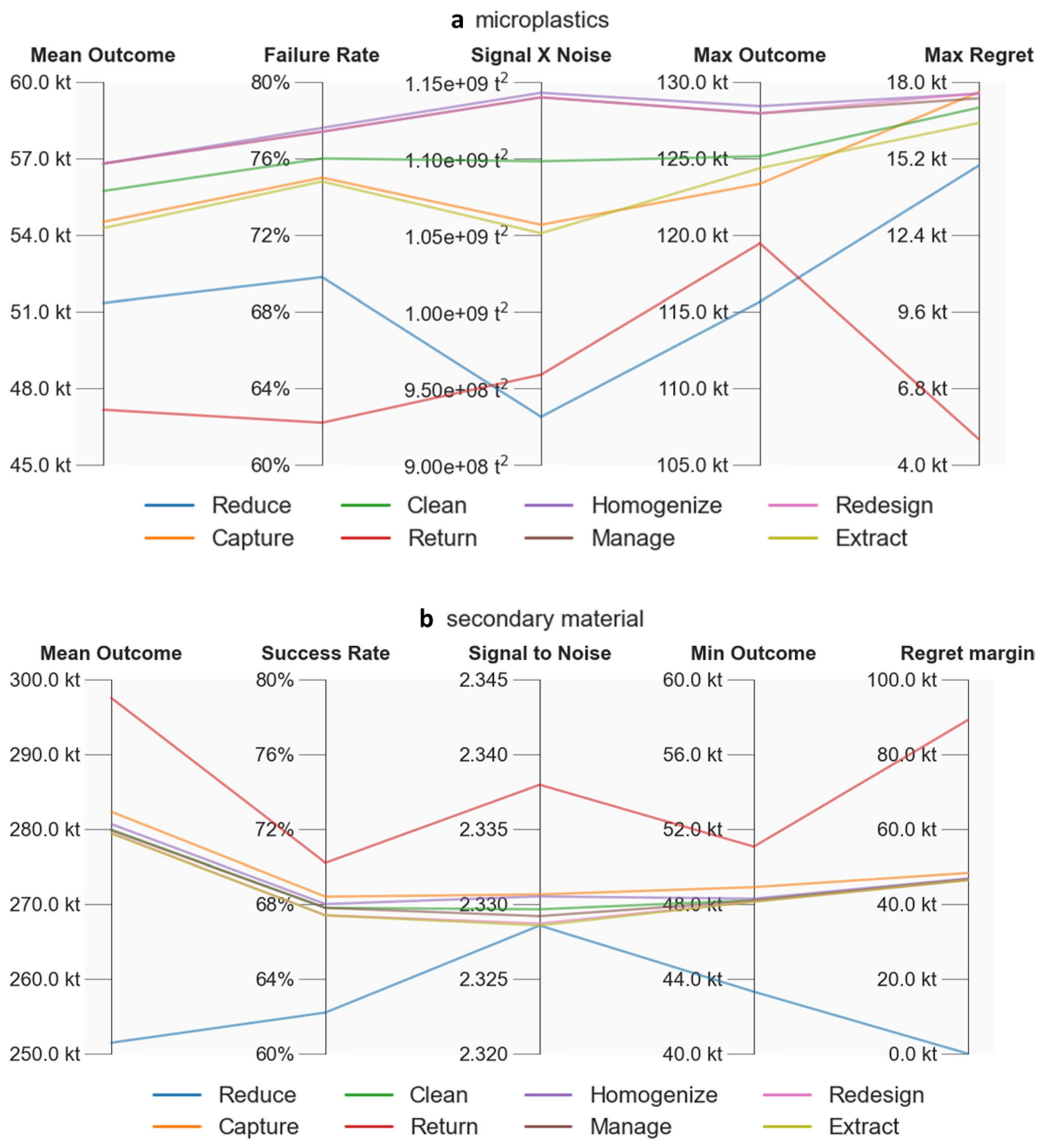

Fig. 4 Policy robustness across robustness metrics for the policy objectives of **a** reducing environmental microplastic emissions and **b** increasing secondary material production, related to consumer PET bottles in the Netherlands from 2025 to 2030. The Python script used to generate this figure is provided on https://github.com/nweijenb2001/emfma/blob/main/main.py

incinerated, and lost during packaging recycling are key vulnerabilities, particularly for secondary material production. Additional significant vulnerabilities specific to microplastic emissions include: (1) the location of PET bottle consumption, (2) aerial emissions from on-the-go use, (3) dispersion from indoor air, (4) the application of large-scale, PET-contaminated compost to agricultural soil, (5) uncollected mixed waste entering surface waters,

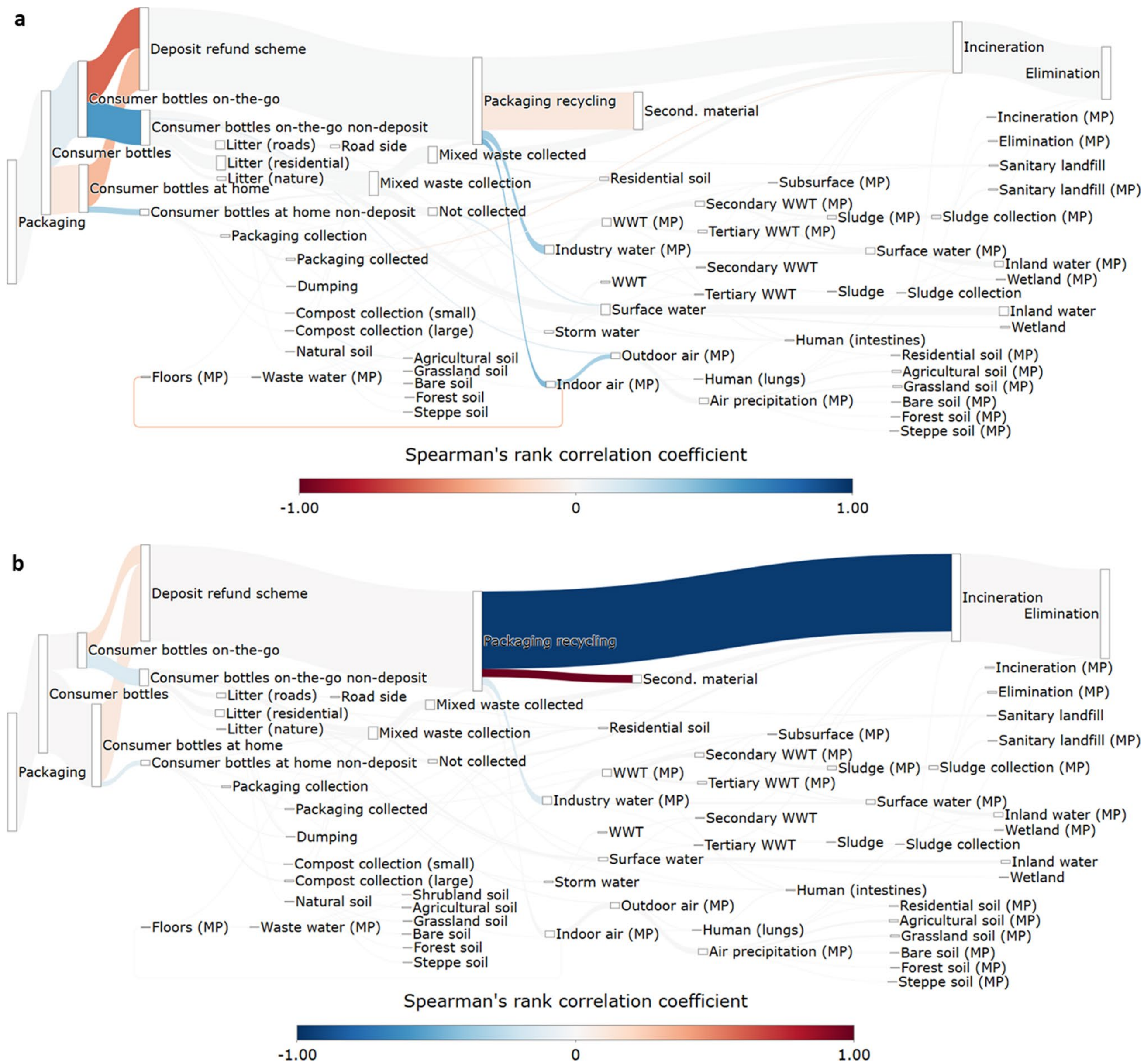

Fig. 5 Policy vulnerability in the business-as-usual case, depicted under the worst-case scenario, for the policy objectives of **a** reducing environmental microplastic emissions and **b** increasing secondary material production, related to consumer PET bottles in the Netherlands from 2025 to 2030. The Python script used to generate this figure is provided on https://github.com/nweijenb2001/emfma/blob/main/main.py

and (6) direct incineration of collected packaging. These robustness controls highlight critical areas where improved data quality could reduce uncertainty and enhance the reliability of policy decisions. However, they do not necessarily represent leverage points for intervention, because system components that appear low-impact under current variability ranges may still become highly influential under policy-driven changes that move parameters outside those ranges.

5 Discussion

The case study illustrated how applying EMA to MFA introduces a distinct conceptual orientation compared to conventional approaches to uncertainty analysis in MFA. The focus shifts to identifying policies that perform reliably across diverse conditions and despite uncertainty instead of identifying the most probable outcomes (Abbasi et al., 2023) or reducing uncertainty (Dong et al., 2023). This shift can support more precautionary and robust decision-making when probabilities are poorly understood. Despite their conceptual differences, the two approaches remain methodologically similar. As in several probabilistic MFA studies, we used expert elicitation for uncertainty characterization (Dong et al., 2023; Mathieux & Brissaud, 2010; Montangero & Belevi, 2007) and applied a form of Markov Chain Monte Carlo sampling for uncertainty propagation (Gottschalk et al., 2010; Lupton & Allwood, 2018). We also conducted sensitivity analysis using Spearman's rank correlation coefficient, following Cao et al. (2018) and Wiedenhofer et al. (2019), to identify robustness controls within the system.

One distinguishing feature of our framework is its use of diverse robustness metrics derived from descriptive statistics of model experiments. While scientific advice often aims for definitive conclusions, multiple perspectives are needed for objective, science-based policymaking (Bschir & Lohse, 2024; Cimorelli & Stahl, 2005; Stirling, 2010). This approach contrasts with probabilistic MFA studies that often report means and standard deviations, implying a certain level of confidence around central estimates (Abbasi et al., 2023; Chen et al., 2023; Kawecki & Nowack, 2019). Under conditions of data scarcity or expert judgement, such confidence may be misleading, as the choice of probability distribution is almost arbitrary (Clavreul et al., 2013) and standard deviations may fail to capture the spread of skewed distributions.

Our study also highlights several practical benefits of combining EMA and MFA. First, it reduces dependence on highly precise input values by relying on plausible uncertainty bounds that do not need to be narrowly defined. This tolerance for imprecision lowers participation barriers in elicitation, particularly for stakeholders who may be unwilling or unable to provide exact estimates due to confidentiality concerns or limited knowledge. Second, applying policies to the same set of scenarios enables direct performance comparisons. Even when policies exhibit similar outcome distributions, grouping results by scenario can reveal consistent relative differences, as observed in the horizontal lines of Fig. 3. Finally, the framework's use of mass-balance-constrained elicitation, combined with two-blocked Gibbs sampling, avoids the need for truncation, normalization of uncertainty distributions (Kawecki & Nowack, 2019), or least-squares fitting of expert judgments (Dong et al., 2023), thereby preserving the integrity of expert judgment.

Nevertheless, combining EMA and MFA has several limitations. First, the applicability of the framework is constrained by the scale of the MFA model. As the number of processes and parameters increases, the elicitation becomes increasingly time-demanding, which can substantially reduce expert participation. In this study, the 117 survey questions already approached the limits of practical feasibility, with contacted experts citing the considerable time burden as the primary reason for non-participation. In addition, the computational burden associated with large-scale models restricts the number of samples that can be feasibly generated and evaluated, further limiting the depth of uncertainty exploration. However, it is important to note that many MFA studies operate at smaller scales, involving fewer than 10 processes, which makes the approach more manageable (Allesch & Brunner, 2015). Secondly, the framework does not account for uncertainty in policy levers or secondary policy effects unless explicitly modeled. For example, it does not capture uncertainty in the magnitude of policy effects, such as the extent to which filter installation at a recycling facility would reduce microplastic release under the Capture policy. It also does not model indirect effects, such as how improved filtration might indirectly alter the amount of plastic sent for recycling. Finally, the reliance on Spearman's rank correlation restricts the detection of higher-order interactions among uncertainties, limiting the identification of robustness controls (Cao et al., 2018; Wiedenhofer et al., 2019).

Various strategies could be employed to address these limitations. For example, the time burden of elicitation can be reduced through an iterative approach in which the analyst defines intentionally wide uncertainty ranges for all external factors and experts are engaged only to refine the ranges of the robustness controls (Fig. 2). Alternatively, AI-based tools may offer potential to support the elicitation process by suggesting initial uncertainty ranges. However, their reliability is likely to depend on the specific MFA domain. Second, the computational burden could be addressed by investigating more efficient sampling strategies that preserve the sum-to-one constraint of transfer coefficients without distorting uncertainty bounds. Third, uncertainty in policy levers can be explicitly incorporated by introducing noise factors or varying lever values, while secondary policy effects can be explicitly discussed and modeled. Finally, isometric log-ratio transformations could be applied to decorrelate transfer coefficients, thereby enabling variance-based sensitivity analysis and the detection of higher-order interactions (Filzmoser et al., 2010). These examples illustrate potential directions for improving the current framework, but alternative approaches may also be possible.

Overall, combining EMA with MFA appears particularly well suited for policy analysis under conditions of data scarcity or deep uncertainty, provided that the model scope remains sufficiently limited to keep both elicitation and computation feasible. When sufficient data is available to construct empirical probability distributions, probabilistic approaches may be more appropriate. Similarly, if policymakers are primarily interested in estimating most likely outcomes rather than ensuring robustness across a wide range of conditions, probabilistic MFA may be the preferred option.

6 Conclusion and future work

The exploratory material flow modeling and analysis framework developed in this study presents a novel approach to policy evaluation in MFA under conditions of deep uncertainty, which is characterized by limited information on probability distributions and a lack of consensus regarding system boundaries, model structures, and evaluation criteria. Unlike probabilistic MFA, which aims to identify the best policy under probable conditions, our framework focuses on identifying the most robust policy under plausible conditions by stress-testing policies across a broad spectrum of scenarios and value systems. The pilot demonstration of the framework showed the effectiveness of the procedures in generating nuanced policy insights that do not overstate the available knowledge nor risk overlooking critical vulnerabilities.

Several aspects of the framework could not be fully tested due to limited expert involvement and the computational burden associated with the large-scale MFA model used in this study. Future research could strengthen the framework by exploring strategies to reduce both computational and elicitation demands, incorporating policy uncertainty and secondary policy effects, and exploring methods capable of capturing higher-order interactions in vulnerability analysis. Additional validation of the framework's practical usefulness would benefit from case studies involving deeper stakeholder engagement, the inclusion of material imports and model relationships as sources of uncertainty, and more rigorous quantification of policy levers to improve the robustness of the analysis. Future work could also investigate policy synergies through the combined evaluation of multiple interventions rather than isolated policy measures. Finally, integration of the framework with existing modeling platforms such as the DPMFA package (Kawecki et al., 2021) or the EMA Workbench (Kwakkel, 2017) could support more systematic application in practice.

7 Summary:

Supplementary Information SI1: This supporting information provides (1) additional detail for the methods used in the manuscript and (2) supplementary data for the case study. We describe our novel elicitation procedure for transfer coefficients that accounts for mass-balance constraints, contrast it with the need for post hoc adjustments, and introduce our uncertainty propagation method for preserving elicited uncertainty bounds. We also discuss vulnerability analysis methods considering the mathematical characteristics of MFA models. For the case study, we describe the robustness metrics, modelling simplifications, policy levers, elicitation results, and sensitivity analysis results.

GitHub (https://github.com/nweijenb2001/emfma): This Supporting Information provides (1) Python scripts facilitating the explanation of methods described in Supplementary Information SI1, (2) Excel and CSV files containing case study data and demonstrating the elicitation of uncertainty bounds, (3) Python scripts and modules used to run the case study, and (4) model results and visualizations.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s44498-026-00170-5.

Author contributions N.W. conducted the literature review, developed the framework, performed the coding and analyses, and wrote the manuscript, including all figures. The research problem was conceived by N.T., and the application of exploratory modelling was suggested by W.A. N.T. and W.A. provided regular feedback throughout the project and reviewed the manuscript. A.S. and S.L. developed the MFA model used in the case study, provided regular feedback, and reviewed the manuscript.

Funding This research was supported by TNO (Netherlands Organization for Applied Scientific Research) during the main author's internship.

Data availability The data that supports the findings of this study are available in Supplementary Information S1 and on GitHub (https://github.com/nweijenb2001/emfma).

Declarations

Conflict of interest The authors declare no conflict of interest.

AI statement During the preparation of this work, the main author used ChatGPT [chatgpt.com] to (1) brainstorm suitable sum-constrained and uniform sampling methods for transfer coefficients in MFA, (2) enhance the author's understanding of statistical concepts such as Dirichlet distributions and variance-based sensitivity analysis, (3) help develop and debug Python code used in the illustrative application, and (4) improve the clarity, conciseness, flow, and academic tone of this paper. Additionally, the main author used Elicit [elicit.com] to help search for academic papers on selected topics. All outputs from these tools were critically reviewed and revised by the main author and co-authors.

References

Abbasi, G., Hauser, M., Baldé, C. P., & Bouman, E. A. (2023). A high-resolution dynamic probabilistic material flow analysis of seven plastic polymers; A case study of Norway. *Environment International, 172*, Article 107693. https://doi.org/10.1016/j.envint.2022.107693

Agusdinata, D. B. (2008). *Exploratory modeling and analysis: A promising method to deal with deep uncertainty*. Delft University of Technology

Allesch, A., & Brunner, P. H. (2015). Material flow analysis as a decision support tool for waste management: A literature review. *Journal of Industrial Ecology, 19*(5), 753–764. https://doi.org/10.1111/jiec.12354

Bankes, S., Walker, W. E., & Kwakkel, J. H. (2013). *Exploratory modeling and analysis* (pp. 532–537). https://doi.org/10.1007/978-1-4419-1153-7_314

Brunner, P. H., & Rechberger, H. (2016). *Handbook of material flow analysis*. CRC Press. https://doi.org/10.1201/9781315313450

Bschir, K., & Lohse, S. (2024). Taking pluralism seriously: A new perspective on evidence-based policy. *Science and Public Policy, 51*(3), 553–556. https://doi.org/10.1093/scipol/scad074

Cao, Z., Shen, L., Zhong, S., Liu, L., Kong, H., & Sun, Y. (2018). A probabilistic dynamic material flow analysis model for Chinese urban housing stock. *Journal of Industrial Ecology, 22*(2), 377–391. https://doi.org/10.1111/jiec.12579

Chen, L., Liu, W., Yang, T., & Nowack, B. (2023). Probabilistic material flow analysis of eight commodity plastics in China: Comparison between 2017 and 2020. *Resources, Conservation and Recycling, 191*, Article 106880. https://doi.org/10.1016/j.resconrec.2023.106880

Cimorelli, A. J., & Stahl, C. H. (2005). Tackling the dilemma of the science-policy interface in environmental policy analysis. *Bulletin of Science, Technology & Society, 25*(1), 46–52. https://doi.org/10.1177/0270467604271221

Clavreul, J., Guyonnet, D., Tonini, D., & Christensen, T. H. (2013). Stochastic and epistemic uncertainty propagation in LCA. *The International Journal of Life Cycle Assessment, 18*(7), 1393–1403. https://doi.org/10.1007/s11367-013-0572-6

Dong, J., Liao, J., Huan, X., & Cooper, D. (2023). Expert elicitation and data noise learning for material flow analysis using Bayesian inference. *Journal of Industrial Ecology, 27*(4), 1105–1122. https://doi.org/10.1111/jiec.13399

Filzmoser, P., Hron, K., & Reimann, C. (2010). The bivariate statistical analysis of environmental (compositional) data. *Science of the Total Environment, 408*(19), 4230–4238. https://doi.org/10.1016/j.scitotenv.2010.05.011

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *Journal of Socio-Economics, 40*(1), 35–42. https://doi.org/10.1016/j.socec.2010.10.008

Gottschalk, F., Scholz, R., & Nowack, B. (2010). Probabilistic material flow modeling for assessing the environmental exposure to compounds: Methodology and an application to engineered nano-TiO2 particles. *Environmental Modelling & Software, 25*, 320–332. https://doi.org/10.1016/j.envsoft.2009.08.011

Hendriks, C., Obernosterer, R., Müller, D., Kytzia, S., Baccini, P., & Brunner, P. (2000). Material flow analysis: A tool to support environmental policy decision making. Case-studies on the city of Vienna and the Swiss lowlands. *Local Environment, 5*, 311–328. https://doi.org/10.1080/13549830050134257

Herman, J. D., Reed, P. M., Zeff, H. B., & Characklis, G. W. (2015). How should robustness be defined for water systems planning under change? *Journal of Water Resources Planning and Management*. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000509

Huijbregts, M. A. J., Gilijamse, W., Ragas, A. M. J., & Reijnders, L. (2003). Evaluating uncertainty in environmental life-cycle assessment. A case study comparing two insulation options for a Dutch one-family dwelling. *Environmental Science & Technology, 37*(11), 2600–2608. https://doi.org/10.1021/es020971+

Kawecki, D., & Nowack, B. (2019). Polymer-specific modeling of the environmental emissions of seven commodity plastics as macro- and microplastics. *Environmental Science & Technology, 53*(16), 9664–9676. https://doi.org/10.1021/acs.est.9b02900

Kawecki, D., Scheeder, P. R. W., & Nowack, B. (2018). Probabilistic material flow analysis of seven commodity plastics in Europe. *Environmental Science & Technology, 52*(17), 9874–9888. https://doi.org/10.1021/acs.est.8b01513

Kawecki, D., Wu, Q., Gonçalves, J. S. V, & Nowack, B. (2021). Polymer-specific dynamic probabilistic material flow analysis of seven polymers in Europe from 1950 to 2016. *Resources, Conservation and Recycling*, *173*, 105733. https://doi.org/10.1016/j.resconrec.2021.105733

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kwakkel, J. H. (2017). The exploratory modeling workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. *Environmental Modelling & Software, 96*, 239–250. https://doi.org/10.1016/j.envsoft.2017.06.054

Kwakkel, J. H., Eker, S., & Pruyt, E. (2016). How robust is a robust policy? Comparing alternative robustness metrics for robust decision-making. In *Robustness analysis in decision aiding, optimization, and analytics*. Springer.

Kwakkel, J. H., & Pruyt, E. (2013). Exploratory modeling and analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change, 80*(3), 419–431. https://doi.org/10.1016/j.techfore.2012.10.005

Kwakkel, J. H., & Pruyt, E. (2015). Using system dynamics for grand challenges: The ESDMA approach. *Systems Research and Behavioral Science, 32*(3), 358–375. https://doi.org/10.1002/sres.2225

Laner, D., Feketitsch, J., Rechberger, H., & Fellner, J. (2015a). A novel approach to characterize data uncertainty in material flow analysis and its application to plastics flows in Austria. *Journal of Industrial Ecology, 20*(5), 1050–1063. https://doi.org/10.1111/jiec.12326

Laner, D., & Rechberger, H. (2016). *Material flow analysis* (pp. 293–332). https://doi.org/10.1007/978-94-017-7610-3_7

Laner, D., Rechberger, H., & Astrup, T. (2014). Systematic evaluation of uncertainty in material flow analysis. *Journal of Industrial Ecology, 18*(6), 859–870. https://doi.org/10.1111/jiec.12143

Laner, D., Rechberger, H., & Astrup, T. (2015b). Applying fuzzy and probabilistic uncertainty concepts to the material flow analysis of palladium in Austria. *Journal of Industrial Ecology, 19*(6), 1055–1069. https://doi.org/10.1111/jiec.12235

Lase, I. S., Tonini, D., Caro, D., Albizzati, P. F., Cristóbal, J., Roosen, M., Kusenberg, M., Ragaert, K., Van Geem, K. M., Dewulf, J., & De Meester, S. (2023). How much can chemical recycling contribute to plastic waste recycling in Europe? An assessment using material flow analysis modeling. *Resources, Conservation and Recycling, 192*, Article 106916. https://doi.org/10.1016/j.resconrec.2023.106916

Lempert, R., Popper, S., & Bankes, S. (2003). *Shaping the next one hundred years: New methods for quantitative, long-term policy analysis*. RAND Corporation. https://doi.org/10.7249/MR1626

Liao, J., Deng, S., Huan, X., & Cooper, D. (2025). Bayesian model selection for network discrimination and risk-informed decision-making in material flow analysis. *Journal of Industrial Ecology*. https://doi.org/10.1111/jiec.70034. n/a.

Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal Distributions across the Sciences: Keys and Clues: On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into v. *BioScience, 51*(5), 341–352. https://doi.org/10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2

Lupton, R. C., & Allwood, J. M. (2018). Incremental material flow analysis with Bayesian inference. *Journal of Industrial Ecology, 22*(6), 1352–1364. https://doi.org/10.1111/jiec.12698

Marchau, V. A. W. J., Walker, W. E., Bloemen, P. J. T. M., & Popper, S. W. (2019). *Decision making under deep uncertainty*. Springer International Publishing. https://doi.org/10.1007/978-3-030-05252-2

Mathieux, F., & Brissaud, D. (2010). End-of-life product-specific material flow analysis. Application to aluminum coming from end-of-life commercial vehicles in Europe. *Resources, Conservation and Recycling, 55*(2), 92–105. https://doi.org/10.1016/j.resconrec.2010.07.006

Maxim, L., & van der Sluijs, J. P. (2011). Quality in environmental science for policy: Assessing uncertainty as a component of policy analysis. *Environmental Science & Policy, 14*(4), 482–492. https://doi.org/10.1016/j.envsci.2011.01.003

Montangero, A., & Belevi, H. (2007). Assessing nutrient flows in septic tanks by eliciting expert judgement: A promising method in the context of developing countries. *Water Research, 41*(5), 1052–1064. https://doi.org/10.1016/j.watres.2006.10.036

Ng, K. W., Tian, G., & Tang, M. (2011). Dirichlet and related distributions. *Wiley*. https://doi.org/10.1002/9781119995784

Park, T., & Lee, S. (2022). Improving the Gibbs sampler. *Wires Computational Statistics, 14*(2), Article e1546. https://doi.org/10.1002/wics.1546

Pauliuk, S., & Heeren, N. (2021). Material efficiency and its contribution to climate change mitigation in Germany: A deep decarbonization scenario analysis until 2060. *Journal of Industrial Ecology, 25*(2), 479–493. https://doi.org/10.1111/jiec.13091

Peano, L. J., Kounina, A., Magaud, V., Chalumeau, S., Zgola, M., & Boucher, J. (2020). *Plastic leak project methodological guidelines*. Quantis and EA. https://quantis.com/report/the-plastic-leak-project-guidelines/

Quik, J. T. K., Hids, A. R., Steenmeijer, M. A., Mellink, Y., & van Bruggen, A. (2024). Emission of microplastics to water, soil, and air. What can we do about it? *Rijksinstituut Voor Volksgezondheid En Milieu RIVM*. https://doi.org/10.21945/rivm-2024-0106

Rijksoverheid. (2021). *Poster Nieuwe wetgeving voor wegwerpplastics*. Ministerie van Infrastructuur En Waterstaat. https://www.rijksoverheid.nl/documenten/brochures/2021/02/08/nieuwe-wetgeving-voor-wegwerpplastics

Rijksoverheid. (2023). *National Circular Economy Programme | 2023 - 2030*.

Schep, E., Bergsma, G., Bouwman, P., Santen, W. van, Warringa, G., Backes, C., & Boeve, M. N. (2025). Evaluatie Statiegeld. *CE Delft*. https://ce.nl/publicaties/evaluatie-statiegeld/

Schwab, O., & Rechberger, H. (2017). Information content, complexity, and uncertainty in material flow analysis. *Journal of Industrial Ecology, 22*(2), 247–476. https://doi.org/10.1111/jiec.12572

Schwarz, A. E., Lensen, S. M. C., Langeveld, E., Parker, L. A., & Urbanus, J. H. (2023). Plastics in the global environment assessed through material flow analysis, degradation and environmental transportation. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2023.162644

Slotte, P., Pohjalainen, E., Hanski, J., & Kivikytö-Reponen, P. (2025). Effect of life extension strategies on demand and recycling of EV batteries—Material flow analysis of Li and Ni in battery value chain for Finnish EV fleet by 2055. *Resources, Conservation and Recycling, 215*, Article 108081. https://doi.org/10.1016/j.resconrec.2024.108081

Soares, M., Colson, A., Bojke, L., Ghabri, S., Garay, O. U., Felli, J. K., Lee, K., Molsen-David, E., Morales-Napoles, O., Shaffer, V. A., & IJzerman, M. J. (2024). Recommendations on the use of structured expert elicitation protocols for healthcare decision making: A good practices report of an ISPOR task force. *Value in Health, 27*(11), 1469–1478. https://doi.org/10.1016/J.JVAL.2024.07.027

Stirling, A. (2010). Keep it complex. *Nature, 468*(7327), 1029–1031. https://doi.org/10.1038/4681029a

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., & Contributors, S. 1. 0. (2020). *SciPy 1.0: Fundamental algorithms for scientific computing in Python*. Nature Methods. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.spearmanr.html

Wiedenhofer, D., Fishman, T., Lauk, C., Haas, W., & Krausmann, F. (2019). Integrating material stock dynamics into economy-wide material flow accounting: Concepts, modelling, and global application for 1900–2050. *Ecological Economics, 156*, 121–133. https://doi.org/10.1016/j.ecolecon.2018.09.010