Representation, Propagation, and Interpretation of Uncertain Knowledge in Dynamic Probabilistic Material Flow Models

Nikolaus A. Bornhöft[1,2] · Bernd Nowack[1] · Lorenz M. Hilty[1,2]

Abstract

The determination of the environmental concentration of a pollutant is a crucial step in the risk assessment of anthropogenic substances. Dynamic probabilistic material flow analysis (DPMFA) is a method to predict flows of substances to the environment that can be converted into environmental concentrations. In cases where direct quantitative measurements of concentrations are impossible, environmental stocks are predicted by reproducing the flow processes creating these stocks in a mathematical model. Incomplete parameter knowledge is represented in the form of stochastic distributions and propagated through the model using Monte Carlo simulation. This work discusses suitable means for the model design and the representation of system knowledge from several information sources of varying credibility as model parameter distributions, further evaluation of the simulation outcomes using sensitivity analyses, and the impacts of parameter uncertainty on the total uncertainty of the simulation output. Based on a model developed in a case study of carbon nanotubes in Switzerland, the modeling process, the representation and interpretation of the simulation results are described and approaches to sensitivity and uncertainty analyses are demonstrated. Finally, the overall approach is summarized and provided in the form of a set of modelling and evaluation rules for DPMFA studies.

1 Introduction

Assessing environmental flows and concentrations of anthropogenic pollutants is a crucial step in determining emerging ecological risks of these pollutants. Because for many pollutants quantitative measurements are not feasible, material flow analysis (MFA) [1] and environmental fate modeling [2] have been developed to provide indirect means for exposure assessment. Based on the material inflow into a system, i.e., based on data on production of chemicals or materials, their use in particular products and subsequent pathways through the technosphere and the environment, environmental flows and stocks can be estimated and environmental

concentrations over time derived. However, for many pollutants, uncertainties about the underlying transfer and fate processes compromise model reliability and the suitability of the models to predict environmental stocks and concentrations. Scenario analysis has been used to investigate systems under different sets of uncertain assumptions [3–5]. Nonetheless, scenario analysis does not include the assessment of the likelihood of a particular parameter setting.

Bayesian techniques provide methods to explicitly represent uncertain knowledge from various uncertain sources [6]. Diverging assumptions about the (true) value of a model parameter are weighted based on the modeler's degrees of belief and combined into a probabilistic parameter distribution. The results derived from Bayesian models are concluded based on the assumptions and their weighting. In Bayesian networks [7, 8], which are the most widespread Bayesian models, parameters are represented by discrete sets of values and assigned probabilities.

In MFA, uncertainty handling can improve the credibility of a model and open it to a larger range of applications. However, most of the existing methods and tools provide

✉ Bernd Nowack
nowack@empa.ch

1 Empa, Swiss Federal Laboratories for Materials Science and Technology, Lerchenfeldstrasse 5, CH-9014 St. Gallen, Switzerland

2 Department of Informatics, University of Zurich, Binzmühlestrasse 14, CH-8050 Zurich, Switzerland

uncertainty handling only based on simple error propagation or on a limited number of parameter distribution functions, e.g., in Umberto [9] and STAN [10].

Probabilistic material flow analysis (PMFA) [11] has been developed to assess a system of pollutant flows as a steady-state system and to represent and propagate parameter uncertainties using Bayesian modeling techniques. In PMFA, uncertain knowledge is represented using continuous probability distribution functions and the assumptions are propagated with Monte Carlo simulation. In dynamic probabilistic material flow analysis (DPMFA) [12], the static approach of PMFA was extended to consider sequences of consecutive periods and derive absolute stocks based on the periodic flows.

Lupton et al. discuss the approach with a focus on iterative model improvements using Bayesian Inference to update model parameters based on newly available information [13]. DPMFA—as well as other Bayesian approaches—aims to include all plausible assumptions about a system dimension in a parameter distribution.

The main drawback of Bayesian approaches in MFA is, however, the increased modeling effort, i.e., to gather, weigh up, and combine all plausible information about a model parameter. General approaches merging data from several sources under epistemic uncertainty have been discussed in the field of information fusion [14, 15]. While Gaussian approaches require random independent observations to assess an underlying phenomenon using regression methods, prevailing epistemic uncertainty makes those approaches less suitable. Information fusion instead, takes a fuzzy set approach, representing information pieces (e.g., expert statements) the an intervals of possible values and reconciles competing ones into one weighted distribution [16]. Those Bayesian belief functions [17] provide a representation formalism that seems suitable for parameter uncertainty.

The goal of Bayesian modeling approaches is to enable prediction modeling based on best knowledge. However, the specific impacts of the individual assumptions on a simulation result are not directly visible anymore. This is where sensitivity and uncertainty analyses [18] are useful. They determine the relative impact of the model parameters, e.g., a transfer coefficient (TC) of a flow relation, on output variables, e.g., environmental stocks [19]. There are several sensitivity analysis techniques in use [20], of which "direct" differential sensitivities investigate the robustness of the model output variable with regard to a parameter variation. Uncertainty analysis methods such as the sensitivity index and the importance index [21] look at the impact of parameter uncertainty on the uncertainty about an output variable. The importance index ranks parameters based on the share

of the total variance of an output variable that is introduced by the uncertainty range of one parameter. From the specific impact of the model parameters, the most influential ones can be identified and further investigated.

While there is a wide range of methods to model incomplete knowledge and perform sensitivity and uncertainty analyses, there is no specific guideline for DPMFA yet. Sensitivity analyses for probabilistic material flow models were performed ad hoc by decreasing the mean value of a of model parameter by 10% and calculating the resulting relative change of the observed model output variable [11, 22]. Uncertainty analysis was done by multiplying the standard deviation of a parameter distribution with the respective parameter sensitivity. A method was also developed to model both the variability and the uncertainty separately in environmental risk assessment [23].

Exposure assessment of engineered nanomaterials (ENMs) constitutes a good example domain for modeling anthropogenic pollutants. Even though new detection methods for ENM have been under development for some time, e.g., by Mitrano et al. [24], a generic quantitative measurement of environmental concentrations is currently not feasible [25]. Instead, different modeling methods have been applied for the indirect assessment of different nanomaterials [26], such as particle flow analysis [27], MFA [28], and probabilistic MFA [29, 30]. A dynamic model was used to predict the changes of the flows of nanomaterials to the environment over time [31], and DPMFA was used to assess the historic development of the environmental stocks of several ENMs in the European Union [32–34].

In this work, we will discuss the DPMFA modeling and evaluation process in detail and apply it on a case study for assessing environmental stocks of carbon nanotubes (CNTs) in Switzerland as a proof of concept. Within that, the focus is set on the representation of uncertain system knowledge and the different types of model parameters and the robustness of modelling decisions. Moreover, the characteristics of a sensitivity analysis for DPMFA models are discussed, and the impacts on the predicted evolving stocks are analyzed. Finally, based on the impact of the uncertainty range of the particular model parameters and their value ranges, a set of scenarios is developed with the goal to explain the uncertainty of the model output as large as possible with only a small number of assumptions about model parameters. The work concludes with a stepwise procedure for handling uncertainty in DPMFA modelling for environmental exposure assessment—summarizing the parameter choice, uncertainty, and sensitivity analyses. This structured process suggests a guidance for the modeling and evaluation process of future environmental assessment modeling studies that had been yet lacking.

2 Materials and Methods

2.1 CNT Case Study Model

The significant flow processes of CNT through the technosphere into the environment are represented as a DPMFA model [12]. This model consists of flow compartments, stocks, sinks and external inflows. Based on the interplay of these compartments over time, the local material accumulations can be derived. The flows between the compartments are determined by local transfer coefficients (TCs) that define the flow from one compartment to another as a rate of its total outflow. This system of local flow dependencies distributes the inflows entering it. The time-dynamic behavior of the system is represented over a set of discrete, subsequent periods (i.e., years). For each period, system inflows are determined and the resulting internal flows and changes in stocks calculated. Moreover, delay functions define the residence time of the material in stocks and the subsequent release rates.

Incomplete knowledge about the actual values of system inflows and transfer coefficients is represented in the form of Bayesian probability distributions assigning non-zero relative likelihoods to all plausible parameter values. The dependent model output variables are calculated based on these input distributions with Monte Carlo simulation.

The model is implemented using the DPMFA simulation package described in [35]. The package provides a ready-to-use simulation infrastructure to perform Monte Carlo simulation experiments and to evaluate a model for a given set of parameter distributions. It also provides a set of white box components for creating a model by implementing and assembling a specific system behavior. The parameter distributions can be defined either by selecting among mathematical distribution functions or by providing samples.

In the case study, the system investigated to illustrate the DPMFA modeling process covers the material flows of CNTs from the year 2003 to 2020. The object of interest is the material accumulating in the stocks over time, in particular in the two environmental compartments sediment and soil. The evaluation of the stocks is demonstrated for the years 2012 and 2020 to cover both assessments of past values and predictions for the near future. An exemplary in-depth investigation is demonstrated for the predicted sediment stock in the year 2020 by performing uncertainty and sensitivity analyses.

The model structure (Fig. 1) and the subdivision of the system into model compartments and the parametrization of the transfer coefficients are derived from a steady-state

Fig. 1 Schematic and simplified model structure. The production, manufacturing, and consumption including the product in-use stocks are combined into one box, as well as the sewage treatment and the waste incineration processes. The sinks in environmental compartments are represented in dark gray, the technical sinks in light gray. Arrows represent flows between the compartments. A complete description of all model compartments and transfers is provided in the supporting information (SI 1.1—1.4)
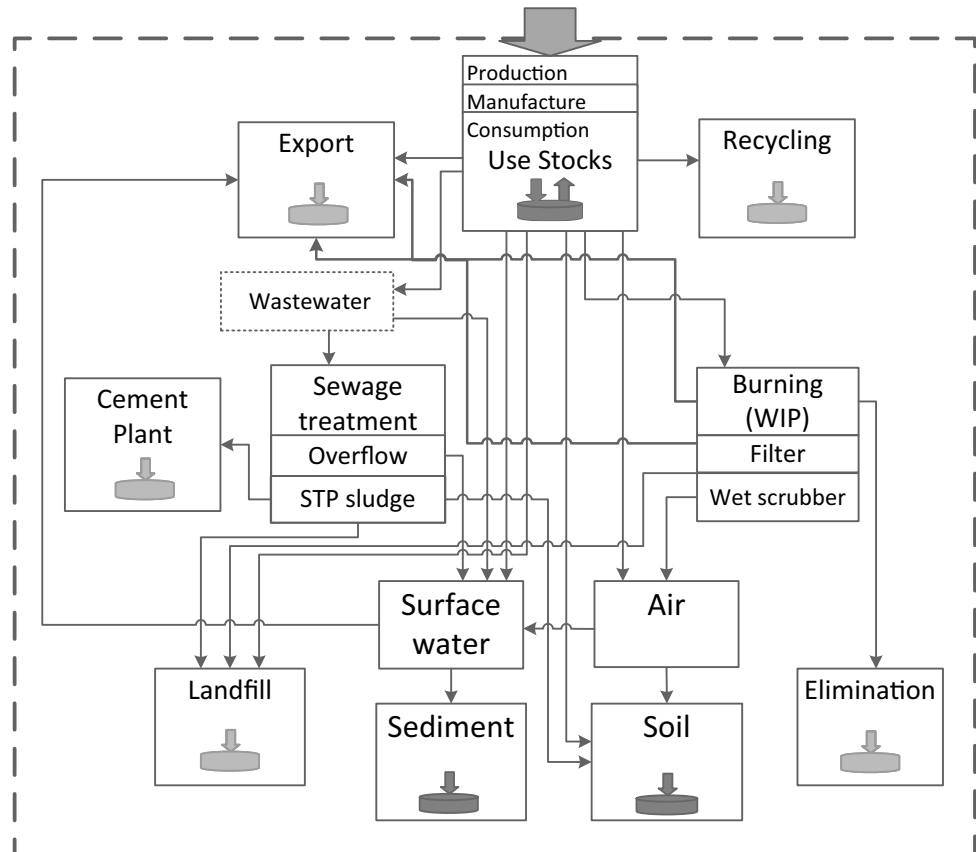
model by Sun et al. [29] and have been used as starting point for the dynamic model described in Bornhöft et al. [12]. The model includes the production of the CNTs, the manufacturing of products containing CNTs, technical processes such as sewage and waste treatment and the receiving environmental media. For the present study, this model was extended by in-use stocks that represent CNTs bound products during the use-phases of the products they are applied in. There are three product categories considered in the model—polymer composites, consumer electronics and automotive to represent the dynamic system behavior. Assumptions about production volumes of CNT are gathered for 2012 from various sources and scaled based on Piccinno et al. [36], where historical production volumes are provided. In total, the model consists of 31 compartments, including the 3 in-use stocks, 7 sinks, and 58 transfer coefficients.

2.2 Uncertainty Representation and Evaluation

Incomplete knowledge about the transfer coefficients and the annual production volumes is represented in the form of Bayesian parameter distributions. This way, epistemic uncertainties about the investigated processes are explicitly represented in the model allowing to assess the reliabilities and the likelihoods of the findings of the simulation study.

The choice of suitable distributions combining information from different sources of varying credibility and ways of representation is based on concepts of information fusion [15, 37]. It describes the merge of diverging information of a common matter to be represented in a single probability distribution. We will describe the transfer of these principles to DPMFA in the following paragraphs.

The robustness of the model regarding different modeling decisions and handling of incomplete knowledge is investigated for (i) the implicit uncertainty range that is added to values originating from data sources that do *not explicitly* provide information about uncertainty and for (ii) the explicit weighting of data from sources of different credibility. For both aspects, variants of the basic model are investigated. To assess the respective contribution of the model parameters to the output variables, direct differential sensitivity analysis is applied. As a deterministic method, it eliminates stochastic influences on the simulation outcome. This analysis is therefore not done with the given stochastic model, but with a deterministic counterpart created by using the parameter distributions' mean values. Sensitivity coefficients:

$$c = \frac{\delta y}{\delta x} \tag{1}$$

describe the impact of a marginal change δx of a parameter x of the flow model on the corresponding change δy

of a model output variable y as coefficient c. Based on the different parameter types of DPMFA, the applicability of differential sensitivity analysis is discussed and applied. To remove the influence of the probabilistic aspect of the model, here the sensitivity is calculated based for a deterministic flow model based using the mean values of the underlying distributions as parameter values.

The absolute influence of the uncertainties in the model parameters on the range of resulting output values is calculated as the difference between the mean values of the investigated dependent model output variable y for the minimum value $x(min)$ and the maximum value $x(max)$ of the model parameter distribution x:

$$dependentRange(y) = abs(y_{x(min)} - y_{x(max)}) \tag{2}$$

The relative uncertainty range (Eq. 2) regards the dependent output range in relation to the mean value of the output distribution as the most likely prediction.

To identify the origin of the uncertainty of the model output y, the dependent uncertainty ranges for all parameter distributions $x_1 \ldots x_n$ are determined. By ranking the parameters according to their contribution to the variables' uncertainties, the most important ones are determined.

Based on the parameters that introduce the largest uncertainties, scenarios are developed. The scenarios aim to reduce most of the model uncertainties to a few assumptions and make their impact explicitly visible. Therefore, instead of using the investigated parameter distribution to simulate the model, a high, a low, and an average deterministic value, each out of the distribution, are used. The 0.05 and the 0.95 quantiles are applied as high and low values. In a subsequent step, the scenarios are combined to investigate the combination of assumptions.

3 Method Application and Results

This section demonstrates the modeling, simulation, and evaluation process along an example application provided by the case study. It focusses on the choice of model parameters for the different system input variables, the interpretation of the model output, particular modeling decisions and their robustness. Moreover, sensitivity and uncertainty analyses are discussed. Along these steps, the procedure, observed results, and their inherent implications are explained in detail.

3.1 Model Design and Simulation

The DPMFA modeling process aims to support the model builder in representing incomplete system knowledge regarding external material inflows, internal transfers, and delay processes as model parameters explicitly. The model

builder is intended to represent the uncertainty of parameters in a realistic and comprehensive way. Based on the uncertain parameters, the dependent model variables for the environmental sinks are calculated and the robustness of some general assumptions is investigated in the simulation process.

3.1.1 External Inflows

External inflows to the model are defined as absolute volumes. Uncertainty is represented in the form of parameter distributions. In the CNT case study, the annual material production was defined as the source. However, data about actual production volumes is sparse. In particular for the years further ago, there are only isolated values for some of the periods. In contrast about the more recent past, more data sources are available. Therefore, from the sources about the recent periods, one comprehensive parameter distribution was developed for the year 2012 as the reference year. Based on a study stating the development of production volumes over time [36] (production volumes of CNTs for Europe and the world, extrapolated to Switzerland based on GDP), scaling factors were defined to adjust the distribution of the reference year to the other years. Scaling factors for missing and future volumes were obtained by extrapolation of the available data (SI 1.6). To generate the parameter distribution for 2012, in a first step, available data sources are gathered and the relevant assumptions worked out, then weighed against each other, and finally merged to a combined distribution. That way, a compromise had to be found between the one-by-one representations of the data sources and a model-wide consistent scheme. The following steps are performed to transform the available data into this form:

- Given likelihood distributions (e.g., observations or samples from previous simulation steps) are used unchanged.
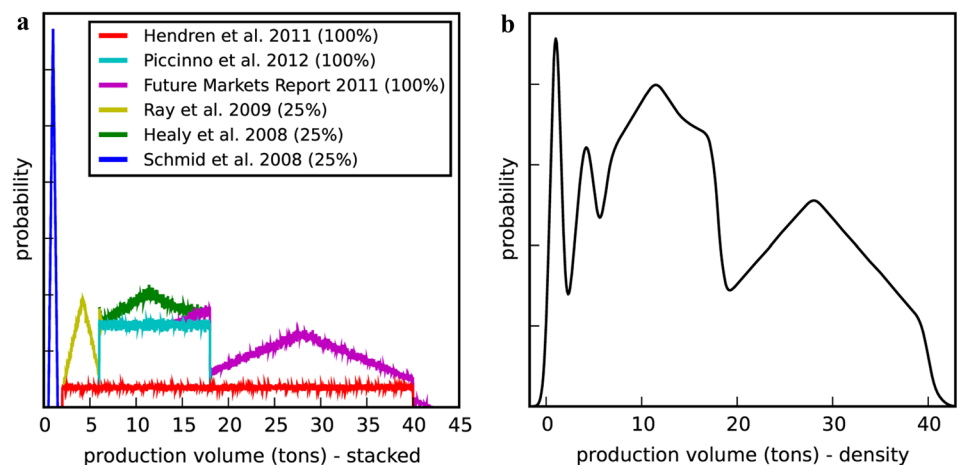- Ranges of plausible values are represented as uniform distributions. This way, all values within the range are represented in the respective model parameter distribution with the same likelihood and without adding further assumptions.
- Single values are represented as triangular distributions, with the value given by the study as the mode value μ and a specific support. The support represents an implicit, plausible value range defining the min and max values of the triangular distribution. This value range includes additional assumptions about the given precision of the value and general considerations about the domain. In the CNT case study a support of ± 0.5 of the mode value μ was used, reflecting the large uncertainties of the domain. This value is based on previous assessments of the uncertainty for this material [38].

Based on the credibility of a data source (e.g., the reliability of the method that was applied in a scientific study or the review process published values have gone through), a relative degree of belief (DoB) is assigned to it. The combined probability distribution of the model parameter is created by merging the single distributions. Depending on the DoB of the data sources, samples of different size are merged into the combined non-parametric distribution to weight their respective impact.

Figure 2 shows the combined parameter distribution of the production volume for 2012 for the case study. Each color displays the share from a particular single distribution, representing the respective weighted share of the data published in one reference. In Fig. 2a, the single distributions are weighted and stacked. In this case, three of the distributions were assumed to have a degree of belief of four times of the other ones and weighted accordingly. The resulting overall distribution of the production volume is shown in Fig. 2b. The combined sample is used as parameter distribution for the CNT flow model. The likelihood distributions of the dependent model variable are inferred from these parameter distributions using a Monte Carlo simulation process.

Fig. 2 Combined belief function for the production volume of CNT in 2012. Diagram (**a**) shows histograms that were sampled from the likelihood distribution of each single study, weighted and added up. Diagram (**b**) shows a density function of the combined sample. Studies published before 2010 were assigned a likelihood of 25% of the newer ones, representing the overall scientific progress in understanding the underlying processes

In a subsequent step, the robustness of the model regarding more similar or more diverse DOBs as well as different supported ranges of plausible values are discussed along the modeling results.

3.1.2 Transfer Coefficients

Parameter distributions for transfer coefficients (TCs) are developed in a manner similar to that of the external inflows as described in the basic articles about the PMFA method [11, 22]. This way, also here Bayesian probability distributions are derived from underlying studies with as little interpretation of the data as possible. However, in this case, a different support is implied for data sources, which only state a single value. As for the absolute system inflows, these values are modeled as triangular distributions with the value referred to in the data source as mode μ. The implicit uncertainty range of a stated value of a TC is based on its minimum distance to 0 or 1. This means that very large and very small TCs are assumed to be less uncertain. For stated TC values ≤ 0.5, the min and max parameter of the triangular distribution are chosen around the stated values as mode value μ and a parameter range from $\mu - 0.5\mu$ and $\mu + 0.5\mu$. For stated TCs > 0.5, a range of $\mu - 0.5(1 - \mu)$ to $\mu + 0.5(1 - \mu)$ is chosen. Analogous to the parameter ranges of the system inflow, the parameter ranges of the transfer coefficients include implicit assumptions about the precision of assumptions made within the domain. More specific knowledge about a transfer would lead to different ranges of implicit plausible values.

For the case study, this is exemplarily illustrated on the CNT removal efficiency of sewage treatment plants (STP) which determines the proportion of the CNTs in the plant that is transferred to STP sludge, i.e., does not remain in the treated water. The parameter distribution is generated from four data sources displayed as diagram in the supplementary information (SI, Fig. 1).

All outgoing TCs from the same model compartment must sum up to 1 to ensure the conservation of mass within the model. Therefore, in each simulation run, one the parameter values are set, the TC of the same origin compartment are normalized to 1.

3.1.3 Delay Times

The development of stocks over time is determined by the material inflows and residence times. The material residence times are parameterized as delay functions that define rates and the time lags after which particular amounts are released from a stock based on the time the material was accumulated. Unlike for other model parameters, the release functions are used to represent delay times are deterministic. In the case study, main delays are determined by the CNTs

being bound during the lifetimes of the products they are used in before they are further released (SI 1.5). As an example of these delay parameters, the residence time of CNTs bound in the "automotive" product category is estimated based on the lifetime distribution of automobiles. Based on a mean value of 11.9 years [43], a normal distribution was used with a distance of 5 years as 1 standard deviation [44]. Figure 3 shows the relative annual release rates computed by a year-wise integration of the distribution function. The annual releases from the stocks contribute to the total model-wide mass flow of CNTs that is calculated for each period.

The discretization of the continuous material releases from stocks into periods of 1-year length is a simplification. However, it corresponds to the way most data is available, for example annual production volumes, and thus appears to be a suitable assumption.

3.2 Model Output

The output of DPMFA models is calculated in a Monte Carlo simulation process that propagates the inherent likelihoods of the parameter distributions to the dependent variables such as the material stocks at a time. These dependent values are made available as samples, whose distribution reflects the likelihood of particular values for the model variable. For each random set of parameter values, the simulation is run over the entire investigated time frame, calculating the dependent model variables.

Figure 4 exemplarily illustrates a density function of the CNT amount accumulated in sediment for the year 2020. It
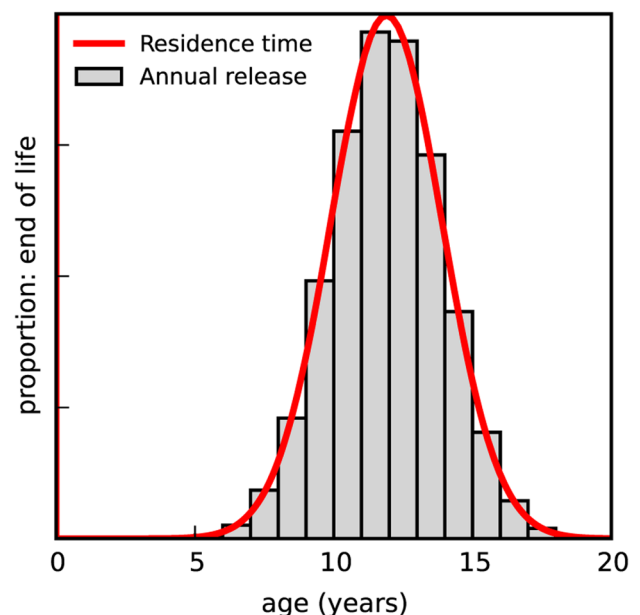
Fig. 3 Relative residence time distribution of CNTs from the "automotive"- in-use stock
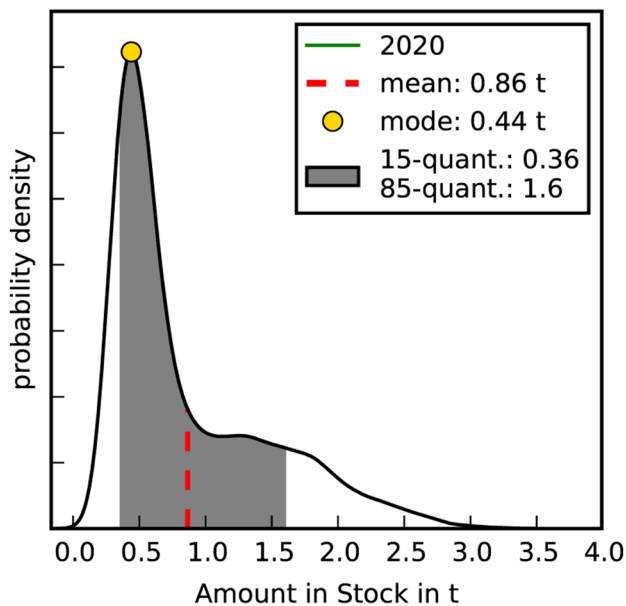
Fig. 4 Projected density function of the CNT sediment stock in 2020

reveals the shape of the function as well as mean and mode values and the 0.15 and 0.85 quantiles.

Table 1 summarizes the model stocks for the years 2012 and 2020 to provide estimations for a year with higher data availability and a forecast for a future period. Uncertainty is indicated here by providing the 0.15 and the 0.85 quantiles in addition to the mean value. The comparison of the stocked amounts shows that a large part of the CNT is bound in product in-use stocks, 48.9 t in 2012 (182 t in 2020). The remaining shares describe amounts that already reached the final model sinks. The material amounts accumulated in environmental media are 0.35 t in soil in 2012 (2.15 t in 2020) and 0.18 t in sediment in 2012 (0.86 t in 2020). The proportion of the material eliminated by waste incineration by that time is much larger, 6.91 t in 2012 (70.3 t in 2020).

Large amounts also end up in technical compartments, especially 12.7 t (63.75) in recycling. The further fate during recycling was not considered in this work, but a first model is available describing the flows out of recycling for selected product categories and nanomaterials [45]. An aggregated overview of all flows summed up over the investigated time frame of 2003 to 2020 is found in the supplementary information (SI, Fig. 2).

3.3 Robustness of Modeling Decisions

As models are representations of a system, idealized for a specific purpose, good modeling decisions focus on the aspects that are decisive for the system behavior under study while abstracting from others to reduce model complexity. The robustness of the simulation results with regard to the modeling and data handling decisions can be used to estimate if a modeling decision taken has considerable impact on the observed model outcome. This robustness reveals aspects of the model for which a more detailed representation of the investigated system could improve the model most to make it more realistic.

In the following, the robustness of the case study model is investigated regarding the data handling decisions made. In the original model, two types of data sources with different credibility stating values for the annual production volume are considered. The type that is considered more credible is weighted four times as strong as the other type.

Table 2 shows the impact of modeling alternatives on the predicted sediment stock in 2020. A stronger weighting of the differences, using a DoB of 1/10 for the data sources with less credibility leads to a mean predicted value of 0.94 t, 8% more than with the basic assumption. An equal treatment of all sources, ignoring their different credibility results in a mean predicted stock of 0.64, 26.4% less than the original model.

Table 1 Model stocks in t for the years 2012 and 2020 with mean values and .15 and .85 quantiles: In-use stocks, elimination, environmental media, and technical compartments		2012			2020		
		.15-quant	Mean	.85-quant	.15-quant	Mean	.85-quant
		% of mean	t/year	% of mean	% of mean	t/year	% of mean
	Composites	68	41.1	133	71	152	129
	Electronics	36	5.33	166	36	19.9	167
	Automotive	34	2.50	170	36	10.6	169
	Elimination	75	6.91	125	78	70.3	122
	Soil	74	0.35	129	78	2.15	122
	Sediment	39	0.18	183	43	0.87	186
	Cement Plt	33	0.06	167	41	0.29	148
	Recycling	55	12.7	146	60	63.8	141
	Export	47	3.12	155	54	16.17	148
	Sum	44.9	72.2	101	223	336	451

Table 2 Alternative modeling of the annual CNT production volumes: mean value and .15 and .85 quantiles for the predicted CNT-Stock in sediment in 2020 weighting the less credible data sources with a 10% degree of belief (DoB) of the more credible ones (row 1) and an equal DoB of all data sources (row 3)

	.15-quant	Mean	.85-quant
DoB of more credible data sources 10× the lower ones	0.40	0.94	1.72
Original model	0.37	0.87	1.62
Same DoB for all data sources	0.26	0.64	1.17

For the second modeling decision, we take a closer look at the assumption of the implicit support for values which are based on a single data source. Table 3 compares the parameter setting of the original model with an increased and a reduced uncertainty range, each by 50% of the original range. Changes of the original assumption of an uncertainty range by increasing or reducing it by 50% of the original range only lead to small changes of the resulting sediment stock in 2020 of 0.01t (<2%). This indicates a high model robustness regarding changes of the implicit support for parameter distributions of the TCs.

3.4 Sensitivity Analysis

Sensitivity analyses investigate the impact of a model parameter on an examined output variable. They allow identifying critical spots of the underlying system that can be addressed in actions to improve the system behavior. For the case study system, these are the processes that affect the development of the environmental CNT stocks. If the predicted environmental concentrations constitute a risk, these spots might be addressed to reduce the environmental exposure [46].

DPMFA models include model parameters and variables of different dimensions. To allow a comparability, relative parameter changes based on differential sensitivity analyzes are investigated. Moreover, the specific characteristics of the different parameter types need to be considered to determine how and to what extent they are suitable for sensitivity analyses.

Table 3 Alternative implicit uncertainty ranges for TCs: Predicted CNT-stock in sediment in 2020 providing the mean value and the .15 and .85 quantile

	.15-quant	Mean	.85-quant
Smaller implicit uncertainty range: $\mu \pm 0.25\,\mu$	0.37	0.86	1.59
Original model	0.37	0.87	1.62
Larger implicit uncertainty range: $\mu \pm 0.75\,\mu$	0.36	0.88	1.60

3.4.1 Inflows, Transfers, and Delays

System inflows are modeled for each year as stochastic likelihood distributions from a continuous value domain. They represent an absolute material inflow for each year. The independent parametrization of external system inflows for each individual period allows considering either a variation of the material inflow of only a single period, or of all periods. As the main objects of interest in the example study are the accumulated stocks, the sensitivity was analyzed with regard to a change of the inflow parameter over all periods using a common growth factor. However, for a closer examination, also other combinations, e.g., a variation of only future periods, could be investigated.

The sensitivity of the model stocks to a variation of the transfer coefficients shows the contribution of that transfer to the development of a stock. This can serve as an indication to find processes within the technosphere, where improvements could reduce the development of environmental stocks. A special characteristic of the method is the assumption of balanced mass flows. If one TC is changed, the TCs of the other flows coming out of the same compartments are normalized to maintain a mass-consistent system behavior. As the consequence of the increase of one flow, the remaining flows are decreased by the same amount. Negative correlations between TCs and stocks are determined by the consideration of mass conservation in the model (e.g., through normalization or direct dependencies). As the sum of all outgoing TCs from a compartment needs to be one, the assumption of an altered TC also implies the adjustment of other, dependent ones. However, to obtain the impact of TCs with several corresponding flows, it is more useful to regard the direct, positive correlations.

While parameters defining material amounts and transfer coefficients take values from a continuous domain (and may be varied by a particular rate), the time representation in DPMFA is discrete, which implies that delay parameters can only be varied in whole time periods. Therefore, a real differential sensitivity analysis to assess the influence of delay times is not possible. However, the overall impact of delay in a temporary stock can be estimated by increasing the delay time by one period and by calculating the model without any delay. An increased delay time of CNT bound in the composite materials in-use stock of the case study would lead to a reduction in the 2020 sediment stock by 0.68%. Assuming an immediate release from composites leads to an increase of the sediment stock by 7.53%.

3.4.2 Sensitivity Analyzes of the Case Study Model

For the case study, a direct differential sensitivity analysis is performed varying all model parameters individually

Table 4 Sensitivity coefficients: impact of relative changes of the model parameters on the material amount in sediment stock in the years 2012 and 2020. The displayed values mark the highest positive and negative impacts

	Sediment 2012	Sediment 2020
Annual production volume (system inflow)	1.01	1.00
TC: STP treatment—> surface water	0.65	0.64
TC: system inflow—> production	0.53	0.52
TC: production—> waste water	0.46	0.45
TC: system inflow—> manufacturing	0.25	0.24
…		
TC: air—> soil	− 1.83	− 2.37
TC: composites—> WIP	− 1.17	− 2.85
TC: system inflow—> consumption	− 77.4	− 75.7

and observing all model stocks as listed in Table 4. The results are discussed in more detail for the sediment stock in 2020. To allow a comparison of the impacts of the parameter changes, they are displayed as relative values. Table 4 provides the largest positive and negative sensitivity coefficients for the case study model regarding the investigated sediments stocks in 2012 and 2020. A table of the impact of each model parameters on the output variables of the environmental stocks for 2012 and 2020 is found in the supporting information (SI Table 4).

The strongest positive impact is found for the annual production volumes with 1.00 for 2020 (1.01 for 2012). This reflects the fact that the model only includes one external source from which all CNTs later accumulated in stocks originate from. It is followed by the sensitivity coefficient of the TC of the flow rate from the STPs compartment to the surface water compartment 0.64 (0.65) and the one for the rate being lost in the production process of 0.52 (0.53). Here, parameter changes have the largest influence on the sediment stock as model output variable. Hence, improving the parameter TC STP treatment—> surface water is most likely to improve the overall system most.

3.5 Uncertainty Analysis

Uncertainty analysis is applied to determine the origin of the uncertainty about a model variable. For the case study, the impacts of the particular parameter distributions on the total uncertainty about the predicted sediment stock in 2020 were investigated in detail. For each model parameter, the model was simulated using the smallest and the largest value of the parameter distribution as a deterministic parameter value. The remaining parameters were kept unchanged. Figure 5 shows the resulting ranges of the predicted mean values. The precise values are listed in the SI (SI Table 5).

The largest influence comes from the annual production, where the smallest volumes that are considered plausible lead to a most likely stock of 0.03 t and the highest one of 2.07 t, representing a range of 2.04 t. Referring to the predicted mean stock from the basic model of 0.85 t, this range is 240% of the most likely assumption. The sewage treatment plant (STP) efficiency introduces an uncertainty range of 1.63 t (191%) followed by the TC of the allocation from material consumption to paints 0.47 t

Fig. 5 Impact of the model parameter ranges on the predicted environmental stock of CNTs in sediment. The bars describe the range from the expected mean in sediment from the minimum to the maximum value of the range of the parameter distributions
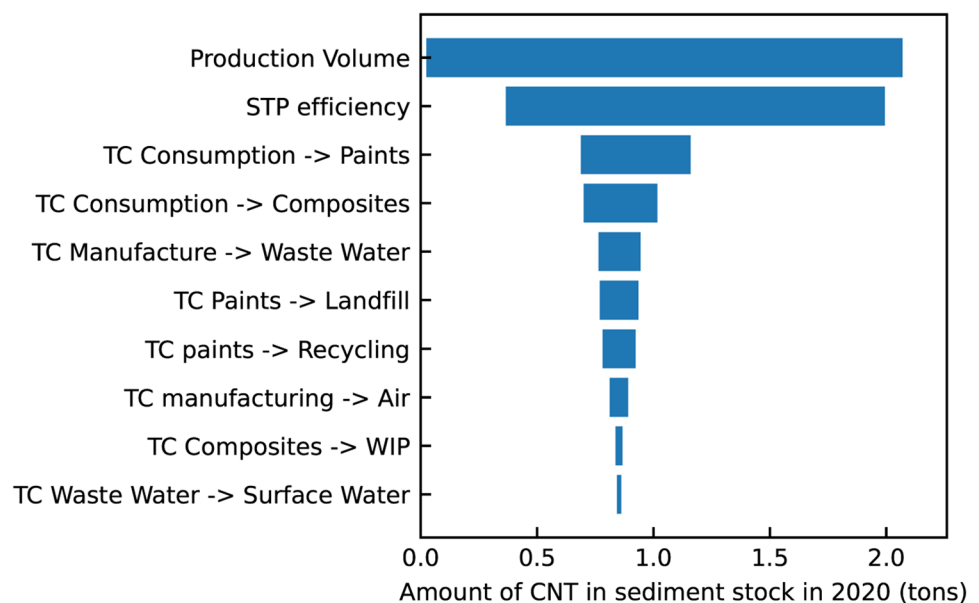
Table 5 Objectives and implementation of the modeling and evaluation steps in DPMFA studies

	Aim/research subject	Application in DPMFA
Model development and simulation	Prediction of stocks of the investigated substance and the environmental exposure	Identification of decisive stock and flow processes
		Representation of all available system knowledge as parameter distributions
		Application of a clear and transparent standardized modeling process
		Robustness checks for important design decisions to determine their impact on the model outcome
Sensitivity analysis	Drivers of the emerging environmental concentrations	Calculation of sensitivity coefficients between the model parameters and the investigated output variable
	Entry points for improvements measures	Parameter mean values as basis
Uncertainty analysis	Contribution of the parameter uncertainty to the overall uncertainty of an output variable	Calculation of output ranges of an investigated model variable for each individual parameter between a high and a low quantile of the parameter distribution
	Identification of points, where better data can improve the model most.	Consideration of scenarios combining the assumptions from the parameters with the highest uncertainty contribution
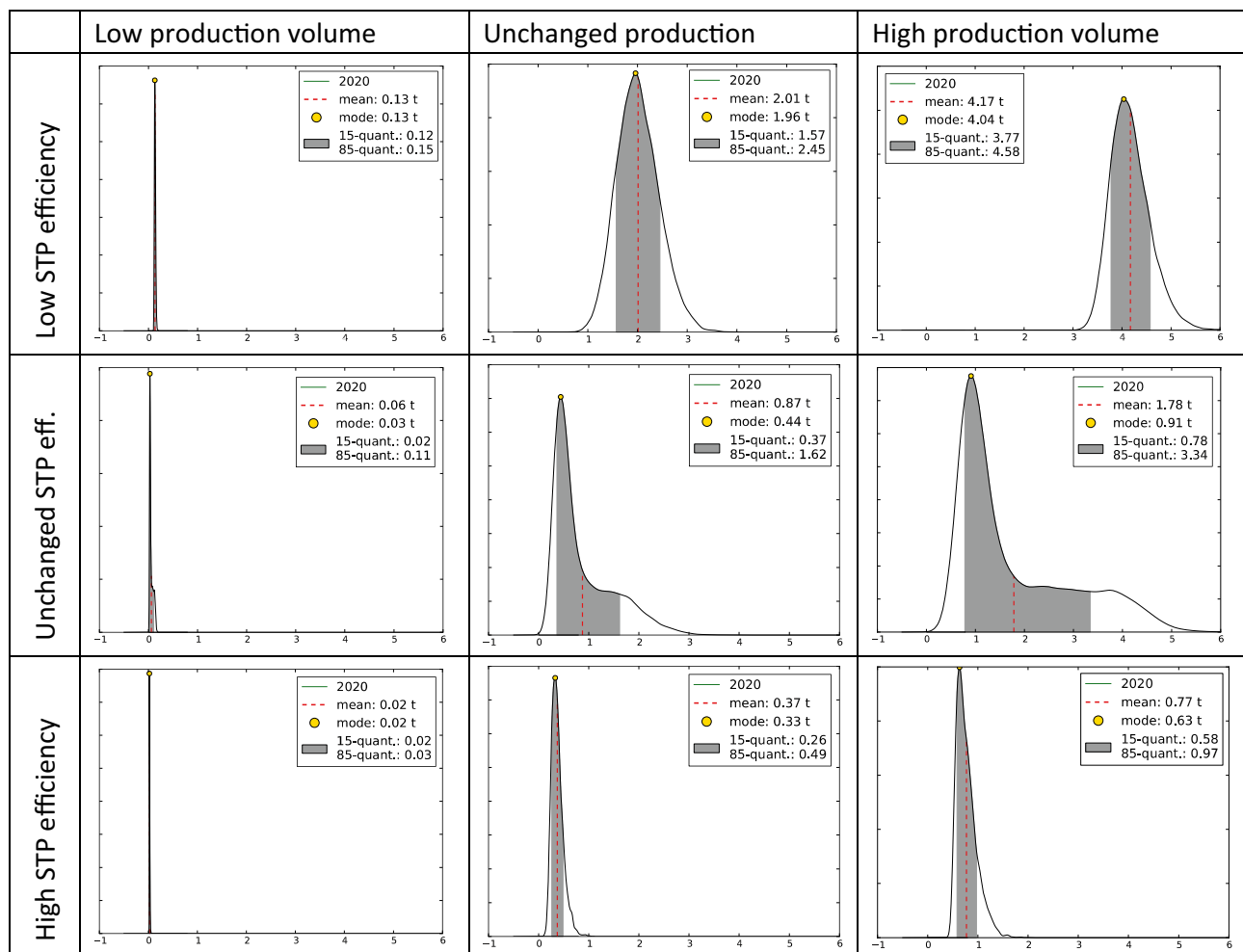
Fig. 6 Scenarios investigating high and low production volumes and STP efficiencies; for the low STP efficiency scenario the .05 percentile and for the high STP efficiency scenario the .95 percentile from the respective distribution are used. The production scenarios use the .05 and the .95 percentile from the production distributions of every year. The other parameter distributions are left unchanged. For each scenario the probability density function of the sediment stock in 2020, its mean, and mode, as well as the .15 and .85 percentiles, are given

(55%), consumption to polymer composites 0.32 t (37%), and manufacturing to waste water 0.18 t (21%).

After uncertainty analysis on a basis of single parameters, the impact of combinations of particular assumptions can be investigated. Therefore, for the model parameters with the largest uncertainty contribution to the investigated model variable, both a low and a high assumption are considered. For the CNT flow model, the uncertainty about the predicted emerging sediment stock in 2020 is strongly determined by the uncertainty about the true material production volume and the STP efficiency. Hence, these parameters are investigated more in detail.

From both parameter distributions, the 0.5 and the 0.95 percentile are taken as plausible low and high assumptions. These assumptions and combinations of them are investigated as different scenarios (Fig. 6). The high production scenario leads to double the amount of CNT in the sediment stock for 2020 of the basic model, 1.78 t and also a much broader uncertainty range. A low STP efficiency would most likely lead to sediment stock of 2.01 t while the combination of a high production and a low STP efficiency results in a mean prediction of 4.17 t–479% of the basic prediction. A low production volume leads to a strongly reduced predicted stock of 6.9% (2.3% in the high STP and 19.9% in the low STP efficiency scenario).

4 Discussion

The proposed procedure for model design, sensitivity, and uncertainty analysis specifies a series of concrete modeling and evaluation steps for predictive modeling of environmental concentrations of anthropogenic pollutants using DPMFA. This way it makes the modeling process transparent and helps to assess the obtained results. Moreover, the CNT case study provides a comprehensive hands-on illustration as an example for a long-lasting anthropogenic pollutant. Table 5 summarizes the main modelling and evaluation steps.

The DPMFA approach uses Bayesian knowledge representation to reproduce epistemic system uncertainties. It allows predicting environmental stocks including the inherent uncertainties. However, as a drawback, it raises the overall modeling effort and the need to explain the obtained results compared to deterministic approaches. Standardized steps from information fusion formalize and streamline the shaping of parameter distributions and help to cope with the rising complexity. Nevertheless, they also allow introducing more complex parameter distributions where existing system knowledge requires it.

Sensitivity analyses identify the main drivers of a dependent system variable. They can serve as a preselection of entry points for measures to reduce environmental stocks

and concentrations. Besides a general reduction of the material production, the improvement of sewage treatment and the reduction of losses during production processes have been identified to affect the resulting stocks of the case study model most. However, as the model includes uncertain assumptions, the applied deterministic differential sensitivity analysis focusing on the means of the distributions is subject to these uncertainties. In particular, this needs to be taken into account in cases where the examined parameters include wide value ranges.

While deterministic flow models are validated within a particular precision and may later be falsified, rejected, and replaced, DPMFA models (like all Bayesian models) are designed to include all plausible values to ensure, they cover the true value as well [47]. Improved system knowledge reduces parameter uncertainty—if a system dimension is known with a higher level of certainty, the parameter distribution representing it becomes narrower and the information derived from the model more definite. The impact of these parameter uncertainties on the model output values was determined using uncertainty analysis. For the case study, the production volume of the material and the STP efficiency introduce the largest uncertainty about the predicted sediment stock. An increase of knowledge about these parameters proposes the largest reduction of uncertainty about the sediment stock. Combining high/low scenarios for these two parameters provide quite clear predictions under the given assumptions.

5 Conclusions

While dynamic probabilistic material flow modeling provides a reasonable approach for environmental exposure assessment in cases where large inherent uncertainties impede the use of deterministic approaches, this work provides an advancement of the available toolset. Applying a rule-based, structured modeling process and sensitivity and uncertainty analysis can increase the conclusiveness of a DPMFA study and provide a more complete picture about the investigated system and the derived model. The derived results provide predictions about environmental stocks and related risks, while the sensitivity analyses identify points for measures to reduce the environmental impacts and uncertainty analyses show where further findings could contribute the most to a reduction of uncertainties. This way, dynamic probabilistic material flow analysis can become an even more meaningful tool for environmental risk assessment in future studies and a valuable approach to estimate hazard to ecosystems through anthropogenic pollutants.

References

1. Baccini, P., & Brunner, P. H. (1991). *Metabolism of the anthroposphere*. Berlin: Germany Springer.
2. MacLeod, M., Scheringer, M., McKone, T. E., & Hungerbuhler, K. (2010). The state of multimedia mass-balance modeling in environmental science and decision-making. *Environmental Science & Technology, 44*(22), 8360–8364. https://doi.org/10.1021/es100968w
3. Huss, W. R. (1988). A move toward scenario analysis. *International Journal of Forecasting, 4*(3), 377–388. https://doi.org/10.1016/0169-2070(88)90105-7
4. Bunn, D. W., & Salo, A. A. (1993). Forecasting with scenarios. *European Journal of Operational Research, 68*(3), 291–303. https://doi.org/10.1016/0377-2217(93)90186-Q
5. Erdmann, L., & Hilty, L. M. (2010). Scenario analysis: exploring the macroeconomic impacts of information and communication technologies on greenhouse gas emissions. *Journal of Industrial Ecology, 14*(5), 826–843. https://doi.org/10.1111/j.1530-9290.2010.00277.x
6. Cullen, A. C., & Frey, H. C. (1999). *Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs*. New York: Plenum Press.
7. Pearl, J. Bayesian networks: a model of self-activated memory for evidential reasoning. In *Proceedings of the 7th Conference of the Cognitive Science Society, University of California, Irvine, 1985* (pp. 329–334). citeulike-article-id:1911949.
8. Ahmadi, A., Moridi, A., & Han, D. (2015). Uncertainty assessment in environmental risk through Bayesian networks. *Journal of Environmental Informatics, 25*(1), 46–59, https://doi.org/10.3808/jei.201500294
9. ifu Hamburg GmbH (2014). Umberto (5.6 ed.).
10. TU Vienna, Institute for Water Quality, Resource and Waste Management, (2012). STAN, subSTance flow ANalysis. (2.5 ed.).
11. Gottschalk, F., Scholz, R. W., & Nowack, B. (2010). Probabilistic material flow modeling for assessing the environmental exposure to compounds: Methodology and an application to engineered nano-TiO_2 particles. *Environ. Modeling Software, 25,* 320–332.
12. Bornhöft, N. A., Sun, T. Y., Hilty, L. M., & Nowack, B. (2016). A dynamic probabilistic material flow modeling method. *Environmental Modelling & Software, 76,* 69–80. https://doi.org/10.1016/j.envsoft.2015.11.012
13. 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
14. Dubois, D., & Prade, H. (2004). On the use of aggregation operations in information fusion processes. *Fuzzy Sets and Systems, 142*(1), 143–161. https://doi.org/10.1016/j.fss.2003.10.038
15. Smets, P. (2007). Analyzing the combination of conflicting belief functions. *Inf. Fusion, 8*(4), 387–412. https://doi.org/10.1016/j.inffus.2006.04.003
16. Džubur, N., Sunanta, O., & Laner, D. (2017). A fuzzy set-based approach to data reconciliation in material flow modeling. *Applied Mathematical Modelling, 43*, 464–480. https://doi.org/10.1016/j.apm.2016.11.020
17. Smets, P. (2005). Belief functions on real numbers. *International Journal of Approximate Reasoning, 40*(3), 181–223. https://doi.org/10.1016/j.ijar.2005.04.001
18. Loucks, D. P., Van Beek, E., Stedinger, J. R., Dijkman, J. P. M., & Villars, M. T. (2005). *Water resources systems planning and management: An introduction to methods, models and applications*: UNESCO.
19. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., et al. (2008). Introduction to sensitivity analysis. In *Global sensitivity analysis. The Primer* (pp. 1–51): John Wiley & Sons, Ltd.
20. Hamby, D. M. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment, 32*(2), 135–154. https://doi.org/10.1007/BF00547132
21. Hoffman, F. O., & Gardner, R. H. (1983). Evaluation of uncertainties in environmental radiological assessment models. In J. E. Till, & H. R. Meyer (Eds.), *Radiological assessments: a textbook on environmental dose assessment*. Washington, DC: U.S. Nuclear Regulatory Commission.
22. Gottschalk, F., Sonderer, T., Scholz, R. W., & Nowack, B. (2010). Possibilities and limitations of modeling environmental exposure to engineered nanomaterials by probabilistic material flow analysis. *Environmental Toxicology and Chemistry, 29,* 1036–1048. https://doi.org/10.1002/etc.135
23. Jacobs, R., Meesters, J. A. J., ter Braak, C. J. F., van de Meent, D., & van der Voet, H. (2016). Combining exposure and effect modeling into an integrated probabilistic environmental risk assessment for nanoparticles. *Environmental Toxicology and Chemistry, 35*(12), 2958–2967. https://doi.org/10.1002/etc.3476
24. Mitrano, D. M., Lesher, E. K., Bednar, A., Monserud, J., Higgins, C. P., & Ranville, J. F. (2012). Detecting nanoparticulate silver using single-particle inductively coupled plasma-mass spectrometry. *Environmental Toxicology and Chemistry, 31*(1), 115–121. https://doi.org/10.1002/etc.719
25. von der Kammer, F., Ferguson, P. L., Holden, P. A., Masion, A., Rogers, K. R., Klaine, S. J., et al. (2012). Analysis of engineered nanomaterials in complex matrices (environment and biota): general considerations and conceptual case studies. *Environmental Toxicology and Chemistry, 31*(1), 32–49. https://doi.org/10.1002/Etc.723

26. Nowack, B. (2017). Evaluation of environmental exposure models for engineered nanomaterials in a regulatory context. *NanoImpact, 8*, 38–47. https://doi.org/10.1016/j.impact.2017.06.005

27. Arvidsson, R., Molander, S., & Sandén, B. A. (2012). Particle Flow Analysis. *Journal of Industrial Ecology, 16*(3), 343–351. https://doi.org/10.1111/j.15309290.2011.00429.x

28. Keller, A., McFerran, S., Lazareva, A., & Suh, S. (2013). Global life cycle releases of engineered nanomaterials. *Journal of Nanoparticle Research, 15*(6), 1–17. https://doi.org/10.1007/s11051-013-1692-4

29. Sun, T. Y., Gottschalk, F., Hungerbühler, K., & Nowack, B. (2014). Comprehensive probabilistic modelling of environmental emissions of engineered nanomaterials. *Environmental Pollution, 185*, 69–76. https://doi.org/10.1016/j.envpol.2013.10.004

30. Gottschalk, F., Lassen, C., Kjoelholt, J., Christensen, F., & Nowack, B. (2015). Modeling flows and concentrations of nine engineered nanomaterials in the Danish environment. *International Journal of Environmental Research and Public Health, 12*(5), 5581–5602. https://doi.org/10.3390/ijerph120505581

31. Song, R., Qin, Y., Suh, S., & Keller, A. A. (2017). Dynamic model for the stocks and release flows of engineered nanomaterials. *Environmental Science & Technology, 51*(21), 12424–12433. https://doi.org/10.1021/acs.est.7b01907

32. Sun, T. Y., Mitrano, D. M., Bornhöft, N. A., Scheringer, M., Hungerbühler, K., & Nowack, B. (2017). Envisioning nano release dynamics in a changing world: Using dynamic probabilistic modeling to assess future environmental emissions of engineered nanomaterials. *Environmental Science & Technology, 51*(5), 2854–2863. https://doi.org/10.1021/acs.est.6b05702

33. Wang, Y., & Nowack, B. (2018). Dynamic probabilistic material flow analysis of nano-SiO_2, nano iron oxides, nano-CeO_2, nano-Al_2O_3, and quantum dots in seven European regions. *Environmental Pollution, 235*, 589–601. https://doi.org/10.1016/j.envpol.2018.01.004.

34. Stoudman, N., Nowack, B., & Som, C. (2019). Prospective environmental risk assessment of nanocellulose for Europe. *Environ. Sci. Nano* (6), 2520–2531. https://doi.org/10.1039/c9en00472f

35. Bornhöft, N. A. (2015). DPMFA - A simulation framework for dynamic probablistic material flow analysis (0.92 ed.).

36. Piccinno, F., Gottschalk, F., Seeger, S., & Nowack, B. (2012). Industrial production quantities and uses of ten engineered nanomaterials in Europe and the world. *Journal of Nanoparticle Research, 14*(9), 1–11. https://doi.org/10.1007/s11051-012-1109-9

37. Destercke, S., Dubois, D., & Chojnacki, E. (2009). Possibilistic information fusion using maximal coherent subsets. *Fuzzy Systems, IEEE Transactions on, 17*(1), 79–92. https://doi.org/10.1109/TFUZZ.2008.2005731

38. Hendren, C. O., Mesnard, X., Dröge, J., & Wiesner, M. R. (2011). Estimating production data for five engineered nanomaterials as a basis for exposure assessment. *Environmental Science and Technology, 45*(7), 2562–2569.

39. Ray, P. C., Yu, H. T., & Fu, P. P. (2009). Toxicity and environmental risks of nanomaterials: Challenges and future needs. *Journal of Environmental Science and Health Part C-Environmental Carcinogenesis & Ecotoxicology Reviews, 27*(1), 1–35. https://doi.org/10.1080/10590500802708267

40. Healy, M. L., Dahlben, L. J., & Isaacs, J. A. (2008). Environmental assessment of single-walled carbon nanotube processes. *Journal of Industrial Ecology, 12*(3), 376–393. https://doi.org/10.1111/j.1530-9290.2008.00058.x

41. Schmid, K., & Riediker, M. (2008). Use of nanoparticles in Swiss industry: A targeted survey. *Environmental Science and Technology, 42*(7), 2253–2260.

42. Future Markets (2012). Nanomaterials in plastics and advanced polymers.

43. Kraftfahrt Bundesamt (2003). Pressemitteilung Nr. 30/2003 - Alter Fahrzeugbestand.

44. Restrepo, E. (2015). Personal communication.

45. Caballero-Guzman, A., Sun, T., & Nowack, B. (2015). Flows of engineered nanomaterials through the recycling process in Switzerland. *Waste Management, 36*, 33–43. https://doi.org/10.1016/j.wasman.2014.11.006

46. Coll, C., Notter, D., Gottschalk, F., Sun, T., Som, C., & Nowack, B. (2016). Probabilistic environmental risk assessment of five nanomaterials (nano-TiO_2, nano-Ag, nano-ZnO, CNT, and fullerenes). *Nanotoxicology, 10*(4), 436–444. https://doi.org/10.3109/17435390.2015.1073812

47. Nowack, B., Baalousha, M., Bornhoft, N., Chaudhry, Q., Cornelis, G., Cotterill, J., et al. (2015). Progress towards the validation of modeled environmental concentrations of engineered nanomaterials by analytical measurements. *Environmental Science: Nano, 2*(5), 421–428. https://doi.org/10.1039/C5EN00100E