



HAL
open science

Quantification of Conceptual Model Uncertainty in the Modeling of Wet Deposited Atmospheric Pollutants

Laura Urso, Mouhamadou Moustapha Sy, Marc Andre Gonze, Philipp Hartmann, Martin Steiner

► **To cite this version:**

Laura Urso, Mouhamadou Moustapha Sy, Marc Andre Gonze, Philipp Hartmann, Martin Steiner. Quantification of Conceptual Model Uncertainty in the Modeling of Wet Deposited Atmospheric Pollutants. Risk Analysis, 2022, 10.1111/risa.13807 . hal-03397627

HAL Id: hal-03397627

<https://hal.science/hal-03397627>

Submitted on 22 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Quantification of Conceptual Model Uncertainty in the Modeling of Wet Deposited Atmospheric Pollutants

Laura Urso ^{1,*} Mouhamadou Moustapha Sy ² Marc-André Gonze,³ Philipp Hartmann,¹ and Martin Steiner¹

Conceptual model uncertainty and parameter uncertainty are dominant contributors to the total uncertainty of a radioecological model output. In the present study the focus is on conceptual model uncertainty, which is often not acknowledged. Conceptual model uncertainty is assessed by subtracting from the total uncertainty of the model output the propagated parameter uncertainty, obtained by means of Bayesian inference analysis. The conceptual model uncertainty is quantified for two process-based models, which describe the interception of wet deposited pollutants under equilibrium and kinetic conditions, respectively. The natural variability due the chemical valence of the elements considered is accounted for in both models. Quantitative evidence has been obtained that the conceptual model uncertainty can contribute to the total uncertainty budget of the models for interception of wet deposited pollutants at least as much as, if not more than, parameter uncertainty.

KEY WORDS: Model structure; Bayesian modeling; uncertainty

1. INTRODUCTION

1.1. Environmental Models and Uncertainty of Their Output

Environmental models deliver outputs that are characterized by an inevitable degree of uncertainty.

Uncertainties can be either of epistemic or aleatory nature (Hoffman & Hammonds, 1994), (Salbu, 2016). Generally, an epistemic uncertainty arises from a misunderstanding of the processes involved or from the lack of relevant data. This can

be reduced by more research or empirical effort. Aleatory uncertainty (also called variability) is a statistical uncertainty, for which quantities vary in space and time because of the pure stochastic nature of the processes involved. Such variability is an inherent property of many natural environments, of processes related to the transfer of pollutants in the environment, of animal and human behavior, and so on. When enough reliable data are available, variability can be quantified and properly accounted for, but cannot be reduced. In the field of radioecology, for example, variability of spatial soil properties, the variability of plant characteristics even within same species as well as variability related to the uptake of radionuclides via food in animals and humans is often acknowledged (Simon-Cornu et al., 2015).

Epistemic uncertainty types that contribute to the total uncertainty of a model output are extensively described by many authors (Ascough II, Maier, Ravalico, & Strudley, 2008; EPA, 2009; Petersen, Janssen, & Van der Sluijs, 2003; Refsgaard, van der

¹Federal Office for Radiation Protection (BfS), Oberschleissheim, Germany.

²Federal Institute for Risk Assessment (BfR), Berlin, Germany.

³Institute of Radiation Protection and Nuclear Safety (IRSN), St-Paul-lez-Durance Cedex, France.

*Address correspondence to Laura Urso, Federal Office for Radiation Protection (BfS), Ingolstaedter Landstrasse 1, 85764 Oberschleissheim, Germany; tel.: +49(0)30183332556; lurso@bfs.de.

Sluijs, Højberg, & Vanrolleghem, 2007; Rowe, 1994; Salbu, 2016; Walker et al., 2003).

Model parameters are constant quantities (Kirchner & Steiner, 2008) whereas model input variables are time-dependent quantities. These are usually measured during experimental activities.

Epistemic uncertainties may include:

- parameter uncertainty, which arises from a limited knowledge of the values for the model parameters and the correlations between model parameters;
- conceptual model uncertainty (sometimes also called structural or model uncertainty), which arises from an inadequate or oversimplified model structure that does not represent key characteristics of an ecosystem and key processes;
- input uncertainty or the uncertainty on the values of the input variables, which arises principally from inaccuracies while measuring these alongside the process of concern.
- scenario uncertainty, which arises from an incomplete knowledge of the scenario to be modeled, especially site-specific information and data;
- modeler uncertainty, which arises from the (imperfect) way in which a modeler translates an assessment situation in a model (inadequate interpretation of the assessment situation, selection of an inappropriate model);
- computational/numeric uncertainty, which arises from inaccuracy in numerical solutions.

In numerous environmental disciplines, uncertainty contributions have been not only characterized but also quantified (e.g., Alderman & Stanfill, 2017; Jin, Xu, Zhang, & Singh, 2010; Lindenschmidt, Fleischbein, & Baborowski, 2007; Radwan, Willems, & Berlamont, 2004). Evidence has been brought that in many cases conceptual model uncertainty explains most of the difference between experimental data and model output (e.g., Draper, 1995; Engeland, Xu, & Gottschalk, 2005; Højberg & Refsgaard, 2005; Neuman, Wierenga, & Nicholson, 2003).

Methods to evaluate conceptual model uncertainty are manifold and no standard recipe is available. However, many make use of the assumption that the residual between experimental output and modeled output quantifies the total uncertainty. For example, starting from probabilistic or Bayesian treatment of parameter uncertainty, the other con-

tributions to the overall uncertainty can be obtained with decomposition of mean squared error as in Alderman & Stanfill (2017), Jin et al. (2010), Draper (1995), and Wallach et al. (2017). If several different conceptual models are available multimodel analysis is possible which requires high computational effort as in Symonds & Moussalli (2011), Walsh & Kaiser (2011), McAllister & Kirchner (2002).

In the context of radioecology, conceptual model uncertainty has not been quantified so far. The present study addresses this gap by considering a specific case study involving two process-based models developed for quantifying the process of interception of wet deposited pollutants including radionuclides.

1.2. Conceptual Model Uncertainty in Radioecology

Radioecological models have often a very simple mathematical structure, for example transfer factors or parametric equations (Urso, Hartmann, Diener, Steiner, & Vives i Batlle, 2015), that lacks of causal attributes and that represents in an aggregated form the processes under consideration. For example, conceptual model uncertainty in radioecology plays an important role when (Gonze & Sy, 2016), (Salbu, 2016):

- empirical, highly sensitive parameters are used, for example the so-called concentration ratios and transfer factors. Concentration ratios quantify the transfer from one medium to another one, for example concentration ratio soil–plant, soil–mushroom, soil–cow milk, and so on. These factors can vary over orders of magnitude and are based on measurements of the total activity concentration in bulk samples (Bq kg^{-1} , Bq m^{-2});
- relevant processes are deliberately excluded, for example resuspension. Resuspension (Pröhl, 2003) is the process by which contaminated particulate material in soil or sediments are remobilized into the air or the water column, respectively, by mechanisms such as wind action, sea currents and mechanical disturbances by human activity. Although generally not considered to be a significant pathway of human exposure it may become an important secondary source of airborne exposure for example due to mining activities because of the larger amounts of dust produced and the increased amounts of

radionuclides that can be inhaled or following a nuclear accident;

- modeling of atmospheric dispersion of the radionuclides is too simple and inadequate. For instance, when sites with complex topography are modeled with Gaussian model, which is appropriate for flat sites (Leelőssy, Lagzi, Kovács, & Mészáros, 2018).
- the radioactive particle characteristics are not considered in the characterization of the source term, for example characteristics of radionuclide inventory of a specific environmental compartment (IAEA, 2011);
- it is assumed that the system is in an equilibrium state when this is not the case (i.e., the dynamics of the process are neglected);

Because of the too simple mathematical structure used in many cases, conceptual model uncertainty together with parameter uncertainty can be considered a major contributor to the total uncertainty budget of model output. Objectives of the present study are to apply a methodology for characterizing and quantifying conceptual model uncertainty to the modeling of interception of wet deposited atmospheric pollutants and to distinguish contribution of the conceptual model uncertainty to the total uncertainty budget of the model output from the contribution provided by parameter uncertainty.

2. MATERIALS AND METHODS

2.1. Total Uncertainty of a Model Output

The present study builds upon the analysis and results of (Gonze & Sy, 2016), in which two process-based models for interception of wet deposited pollutants are developed and analyzed in terms of parameter uncertainty and model performance. In (Gonze & Sy, 2016) as well as in the present study, it is assumed that the measure of the total uncertainty of model output is given by the residual between modeled output and the experimentally obtained quantity. Under this assumption, it is also implicitly granted that the data used to carry out quantitative uncertainty analysis is representative of the studied quantity/process. The effect of variability also impacts the residuals and hence differentiation between uncertainty and variability needs to be addressed. In the present study, variability is assumed to be solely induced by the physicochemical characteristics of the pollutant.

Not all the uncertainty contributions listed in Section 1 propagate directly to the model output. For example, scenario uncertainty is usually addressed by a “What if” type of analysis (Refsgaard, Van der Sluijs, Brown, & Van der Keur, 2006). However, parameter uncertainty and conceptual model uncertainty do propagate to the model output. The authors that produced datasets used in the present study did not provide any information about quality of measurement and/or of the sampling strategy and therefore input uncertainty is not accounted for. However, in (Gonze & Sy, 2016) a (not formal) sensitivity analysis based on the mathematical analysis of the models, estimation of the ranges of variation for input quantities and Monte Carlo calculations was carried out and impact of input variables and model parameters on model output was assessed.

2.2. Equilibrium Model and Kinetic Model for Quantifying Wet Interception Process

Wet deposition is the process for which airborne pollutants are entrained by falling rain droplets and deposit onto the ground surface. In the case of a vegetated surface, wet interception is the process by which a fraction of the deposit is intercepted by the aerial biomass (e.g., shoot). The process of wet interception is quantified in terms of wet interception factor f_B , defined as:

$$f_B = \frac{f}{B} \quad (1)$$

where f is the ratio between the amount of pollutant retained by plant and the total amount of pollutant deposited on the soil plant system. The f_B is obtained by normalizing f to the standing biomass B [kg m^{-2} d.w.] in order to account for different growth phases of the plants. f depends upon the plant/foilage characteristics, precipitation characteristics and element characteristics. The amount of pollutant absorbed by the plant will be due to the balance between drainage and absorption mechanisms, water storage capacity and chemical valence of the pollutant. In fact, it is observed that inert particles interact more with leaf surfaces than divalent cations and these are more reactive than monovalent cations (Gonze & Sy, 2016). Anions (e.g. iodine) do not interact with plant surfaces. Two process-based models are presented in detail in (Gonze & Sy, 2016). These have been derived from mass balance equations with explicit parameterization of the hydrological, biological and chemical mechanisms involved.

The equations used in the analysis are the analytical solutions to the system of mass balance equations described in detail in (Gonze & Sy, 2016). The equilibrium model (EM) has been obtained by assuming that the interaction between pollutant and plant foliage is instantaneous and reversible (Gonze & Sy, 2016). The kinetic model (KM), instead, has been obtained by assuming that the interaction between pollutant and plant foliage is irreversible and depends on the kinetics of absorption and drainage processes. The EM and KM equations are provided here for completeness:

- EM:

$$f \approx \begin{cases} 1 - p & \text{if } T \leq T_s \\ \frac{1-p}{T \cdot \lambda} \cdot (1 - (1 - \lambda \cdot T_s) \cdot e^{-\lambda \cdot (T - T_s)}) & \text{if } T > T_s \end{cases} \quad (2)$$

where

$$\lambda = \frac{D_s}{\text{LAI} \cdot (L + \text{CR})} \quad (3)$$

- KM:

$$f \approx \begin{cases} 1 - p & \text{if } T \leq T_s \\ \frac{1-p}{T \cdot \lambda} \cdot \left[\frac{D_s}{\text{LAI} \cdot L} \cdot T_s + \frac{\alpha}{\text{LAI} \cdot L} \cdot T + \left(\frac{D_s}{\alpha + D_s} - \frac{D_s}{\alpha + (1-p) \cdot I} \right) \cdot (1 - e^{-\lambda \cdot (T - T_s)}) \right] & \text{if } T > T_s \end{cases} \quad (4)$$

where

$$\alpha = \frac{J}{K} \cdot \text{LAI} \quad (5)$$

and

$$\lambda = \frac{\alpha + D_s}{\text{LAI} \cdot L} \quad (6)$$

To use Equation (2), (3), (4), (5) and (6)

- the single-sided leaf area index $\text{LAI} [\text{m}^2 \text{m}^{-2}] = \text{SLA} \cdot B$
- the exposure time $T [h] = \frac{H}{I}$
- the saturation time $T_s [h] = \frac{B \cdot \text{SLA} \cdot L}{(1-p) \cdot I}$
- the drainage term at saturation $D_s [\text{mm h}^{-1}] = (1-p) \cdot I - E_0$

need to be calculated starting from the parameters CR, JK, SLA, L and input quantities B , H , I , p , E_0 listed in Table I.

In Table I, the concentration ratio CR (mm) mathematically represents the interaction between

Table I. Parameters and Input Quantities Involved in EM and KM

L (mm)	Specific foliage storage
SLA ($\text{m}^2 \text{kg}^{-1} \text{d.w.}$)	Specific foliage area
B (kg d.w. m^{-2})	Standing biomass density
E_0 (mm h^{-1})	Evaporation rate
H (mm)	Rainfall height at time of deposition
I (mm h^{-1})	Rainfall intensity
p (-)	Throughfall coefficient
CR (mm)	Concentration ratio
J/K (mm h^{-1})	Absorption velocity

pollutant and plant foliage in the EM whereas the absorption velocity J/K (mm h^{-1}) mathematically represents the interaction between pollutant and plant foliage in the KM. The throughfall coefficient p provides the fraction of rainfall that reaches the ground without hitting the plants. The biomass density B is specific to each herbaceous species whereas the rainfall height H and the rainfall intensity I are specific

to the rainfall event considered in the experiment. Different assumptions for the dependence between wet interception factor and chemical valence have been tested to quantify the impact of variability related to chemical valence. In (Gonze & Sy, 2016), it has been shown that both the models perform best when the chemical valence of the element interacting with plant is accounted for. For anionic particles (i.e. which undergo no chemical interaction with the plant, that is $\text{CR} = J/K = 0$) EM and KM reduce to the same equation. For water (H_2O), the models simplify to a form similar to the Horton equation (Horton, 1919).

2.3. Available Data

Data used in the present study consist of 440 observations for wet interception factor f_B obtained mostly with controlled experiments using monovalent cations (Cs), divalent and trivalent cations (Ba, Be, Cd, Cr, Pb, Sr), inert particles (Polystyrene with median aerodynamic diameter of $1 \mu\text{m}$ and

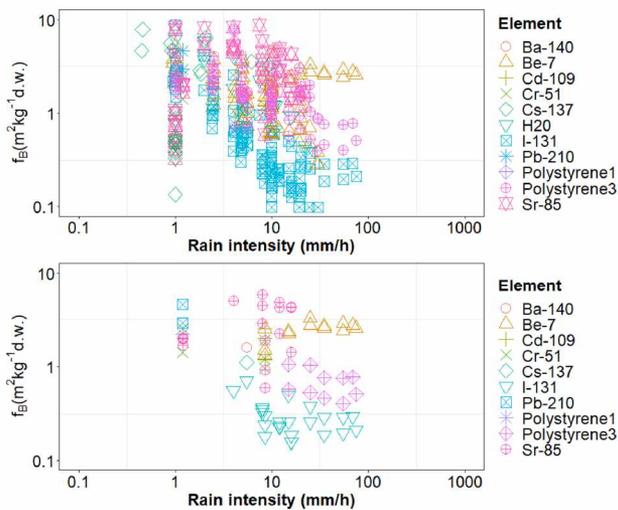


Fig 1. Data used for calibration (top) and prediction (bottom).

3 μm), anions (I), water (H_2O) and different herbaceous species (e.g., grasses, clovers and weeds). Data considered originated from experiments that were conducted either in greenhouses or outdoor plots, in which at least the physicochemical form of the pollutants and the rainfall height were known and documented in the related publications (Chadwick & Chamberlain, 1970), (Angeletti & Levi, 1977b), (Angeletti & Levi, 1977a), (Hoffman, Thiessen, Frank, & Blaylock, 1992), (Hoffman, Thiessen, & Rael, 1995), (Kinnersley, Goddard, Minski, & Shaw, 1997), (Bengtsson, Gärdenäs, Eriksson, Vinichuk, & Rosén, 2014). In addition, three observations from Chernobyl fallout are included (Jacob et al., 1993). The set of collected observations covers a wide range of variation both for rainfall amount and intensity. As an example, in Fig. 1, the experimental wet interception factor f_B is plotted against the rain intensity I (mm h^{-1}). In the experiments considered, observations are obtained with either constant or intermittent precipitation. For an intermittent precipitation, an equivalent rainfall amount and duration were estimated by cumulating event-based quantities, and an equivalent intensity was estimated as the ratio of the cumulated amount and duration.

Of the 440 observations, (Gonze & Sy, 2016) use 363 data points for calibrating the free parameters CR, SLA, L for EM and J/K , SLA, L for KM as shown in Fig. 1 (top). For these datapoints all relevant input variables in the models are available. Data for model calibration consists of 27 datapoints for monovalent cations (Cs), 125 datapoints for di-

valent cations (Be, Sr), 88 datapoints for inert particles (Polystyrene with median aerodynamic diameter of 1 μm and 3 μm), 108 datapoints for anions (I) and 15 datapoints for H_2O . The data for model prediction instead consists of 1 datapoint only for monovalent cations (Cs), 24 datapoints for anions (I), 11 datapoints for inert particles (Polystyrene) and 41 datapoints for divalent and trivalent cations (Ba, Be, Cd, Cr, Pb, Sr). The 77 data points used for prediction analysis are shown in Fig. 1 (bottom). Although the experiments are conducted on different types of herbaceous species, no clear dependency of f_B upon herbaceous species is observed in (Gonze & Sy, 2016).

For the prediction data, information either on the standing biomass B and rainfall height H is not available and is produced “artificially” by means of random sampling over the experimental distribution of H or B of the calibration data.

Although the present study considers the process of wet interception in context of radioecology, not only data for radionuclides are considered but also data available for inert particles (i.e., Polystyrene) and water (H_2O) are used in order to account for the benefit of a larger database to the statistical analysis. The EM and KM, in fact, account for chemical valence of pollutants and do not depend on radioactive properties. Data used in the present study is included in Section 6. Also, it is made extensively available in (Gonze & Sy, 2016).

2.4. Bayesian Inference Analysis and Parameter Uncertainty

In (Gonze & Sy, 2016) Bayesian analysis is carried out for calibrating the model parameters (CR, SLA, L) from EM and (J/K , SLA, L) from KM and for determining their uncertainty. Bayesian inference (Box & Tiao, 2011) consists of calculating the posterior distribution functions for relevant model parameters by means of Monte Carlo Markov Chain simulations on the basis of assumed priors and available observations (stochastic realizations of the modeled process).

The input variables, which are model components for which the value is known are B , H , I . These are driving input quantities specific to each herbaceous plant (for B) or rainfall event (H , I) considered in the laboratory experiment and ideally measured for each realization. Being different among the experiments, they cannot be estimated by any statistical calibration technique.

The model parameters, for which the uncertainty is quantified using the Bayesian analysis are CR, JK, SLA, and L . In particular, SLA and L are driving parameters that moderately vary with herbaceous species and can be assumed constant among the realizations (although uncertain). For the specific foliage storage L , a uniform distribution between 0.1 and 0.5 mm was defined as prior in order to cover the typical range of values found in the literature for the maximum thickness of the water film that can be retained by leaves (Clark, 1940) (Merriam, 1961). For the specific foliage area, SLA the range from 12 to 35 m² kg⁻¹ d.w., consistent with the values in (Jouven, Carrère, & Baumont, 2006), was adopted as such range of values corresponds not only the interception by the foliage but also by the other organs such as the stem. For CR and JK, no preliminary accurate information could be used to appropriately define a prior distribution. However, to avoid implausible values the upper bounds of the uniform distributions were fixed at 10 mm for CR and 100 mm h⁻¹ for JK respectively whereas the lower bounds were fixed to 0 for both quantities to account for nonreactive (i.e., nonabsorbed) species such as iodine.

E_0 and p are quite insensitive quantities and therefore are fixed to a deterministic value. As climatic conditions may have been very specific for especially in-door experiments, it cannot be excluded that relatively high evaporation rates occurred. In (Gonze & Sy, 2016) two extreme situations were considered and tested, by setting E_0 value to either 0 mm h⁻¹ (no evaporation) or 0.6 mm h⁻¹ (high evaporation). The best agreement in (Gonze & Sy, 2016) between model and experimental observations was obtained with $E_0 = 0.6$ mm·h⁻¹. Therefore, in the present study, the evaporation rate is set to $E_0 = 0.6$ mm·h⁻¹. The through-fall coefficient p is set to 0, which is valid for a dense, nonsparse herbaceous layer because the selected experiments were not dealing with sparse vegetation.

The posterior distribution functions of the model parameters (CR, SLA, L) from EM and (JK, SLA, L) from KM are obtained, with which the posterior distribution functions of the models' output f_B is calculated.

In the Bayesian formulation, the quantity f_B is of stochastic nature and is normally distributed (based on the principle of maximum entropy (Jaynes, 1957)) with mean value μ_i from EM and KM output and residual variance S_i^2 obtained for each data point i :

$$f_{B_i} \sim N(\mu_i, S_i^2) \quad (7)$$

In (Gonze & Sy, 2016), 15.000 simulations were used to obtain Bayesian statistics for each of the relevant parameters. The probability distribution functions (pdfs) are obtained for all the relevant parameters and for residual uncertainty S_i .

2.5. Quantification of the Conceptual Model Uncertainty

Given that in the analysis parameter uncertainty and variability are accounted for and given the formalism presented in Section 2.4, it is justified to consider that the residual uncertainty S fully represents the conceptual model uncertainty. Hence, S will be referred to as the conceptual model uncertainty throughout the rest of the study. Predictive calculations for all 77 datapoints ($i = 1, \dots, 77$) are carried out with and without propagating the conceptual model uncertainty S , that is, the posterior distributions of the models' outputs have been calculated from the posterior distributions of the model parameters by using

$$f_{B_i, S_i} = f(p_1, \dots, p_n, S_i) \quad (8)$$

and

$$f_{B_i} = f(p_1, \dots, p_n) \quad (9)$$

The conceptual model uncertainty S is quantified by comparing the two pdfs obtained with Equation (8) and (9), which account for propagated S , \mathbf{p} and only propagated \mathbf{p} , respectively. The pdfs are then compared by means of different statistical metrics in order to evaluate the relative impact of parameter uncertainty and conceptual model uncertainty on the total uncertainty.

2.6. Metrics Used to Compare pdfs and to Quantify Conceptual Model Uncertainty

The conceptual model uncertainty for the EM and KM wet deposition models is quantified by comparing:

- the error bars obtained with the 95th percentile for model output with propagated S and without propagated S ;
- coefficients of variations (Özkaynak, Frey, Burke, & Pinder, 2009).

The error bars for each of the datapoints considered are used to deduce from 95% credibility interval

of model output and 95% credibility interval of parameter uncertainty, the 95% credibility interval of conceptual model uncertainty. From the width of the error bars obtained with and without propagating S , it is possible to visualize how the conceptual model uncertainty S affects the total uncertainty compared to parameter uncertainty for each datapoint.

The coefficients of variations (CV) are calculated using the following mathematical expression:

$$CV_{95} = \frac{f_{B_{P97.5}} - f_{B_{P2.5}}}{f_{B_{\text{median}}}} \quad (10)$$

where $f_{B_{P97.5}}$ and $f_{B_{P2.5}}$ are the f_B values obtained with probability 97.5% and 2.5%, respectively. $f_{B_{\text{median}}}$ is the median value of f_B . The CV_{95} provides the dispersion of the 95th percentile to the median value. The CV_{95} are calculated for EM and KM in order to:

- (1) quantify the difference between the pdf that accounts for the contributions of the conceptual model uncertainty and parameter uncertainty and the pdf that includes only the contribution of the propagated parameter uncertainty;
- (2) quantify the effect of the variability on the model output by comparing f_B obtained by calibrating S over the whole dataset (obtaining a unique range of S for all pollutants, independently of their chemical valence) and f_B obtained by calibrating S over the radionuclide-specific subsets S_i (obtaining single ranges of S for each type of pollutant). The latter is the standard approach in the present study.

3. RESULTS

3.1. Application of Data and Use of Output from Performed Parameter Uncertainty Analysis

Predictive f_B was obtained by using posterior distributions for the parameters inferred in (Gonze & Sy, 2016). Note that, 15,000 values of f_B were obtained by considering the 15,000 values of the posterior distributions of each of the relevant parameters and input values. The posterior distributions of f_B are shown in Fig. 2 for one datapoint. The distribution of f_B were compared to those recovered in (Gonze & Sy, 2016) for each of the datapoints. The values were very similar but not always identical, since for predictive calculations the missing input values were obtained randomly in both studies using different seeds for sampling over the available distributions.

Fig. 2 shows the posterior probability distributions obtained for the model output f_B with and without propagating the conceptual model uncertainty S . The pdfs without S are narrower and account for the contribution of the propagated parameter uncertainty only. The width of the posterior distributions obtained by propagating S is a result of Bayesian inference analysis. The posterior distributions obtained for f_B are not normally distributed.

Data for iodine were used to check the consistency of predictions with EM and KM obtained in (Gonze & Sy, 2016). As mentioned in Section 2.2 for iodine the EM and KM provide identical model outputs. This was confirmed in the present study, indicating also the correct implementation of the models in the programming environment R (Grunsky, 2002) Version 3.6.2.

3.2. Graphical Quantification of Conceptual Model Uncertainty Via Error Bars

Error bars for all 77 datapoints and both KM and EM were produced. An example is displayed in Fig. 3, which shows the difference between EM and KM with and without considering S for obtaining the posterior distributions of f_B for one datapoint for divalent cation (barium and lead), for monovalent cation (cesium) and for anion (iodine). In Section 6 (Supporting Information) the results obtained for 77 datapoints with and without S are presented in tabulated form. The difference between error bars with and without conceptual uncertainty S is in general larger for EM than KM. It is more or less the same in the case of data points for which the two models perform similarly. The difference between error bars with S and without S shows that parameter uncertainty contributes less to the overall uncertainty of the model output than the uncertainty of the model structure. An exception is Cs-137, for which the posterior distribution of f_B without S for KM (due to parameter uncertainty) is already rather large. In fact, the posterior distribution function obtained for J/K for the monovalent cation Cs-137 is larger than for other particles and makes most of the parameter uncertainty.

3.3. Coefficient of Variations (CV_{95})

The conceptual model uncertainty S increases the dispersion of the model output for the KM case on average by a factor 13 (from the ratio of the CVs obtained) and by a factor 22 for the EM case. This

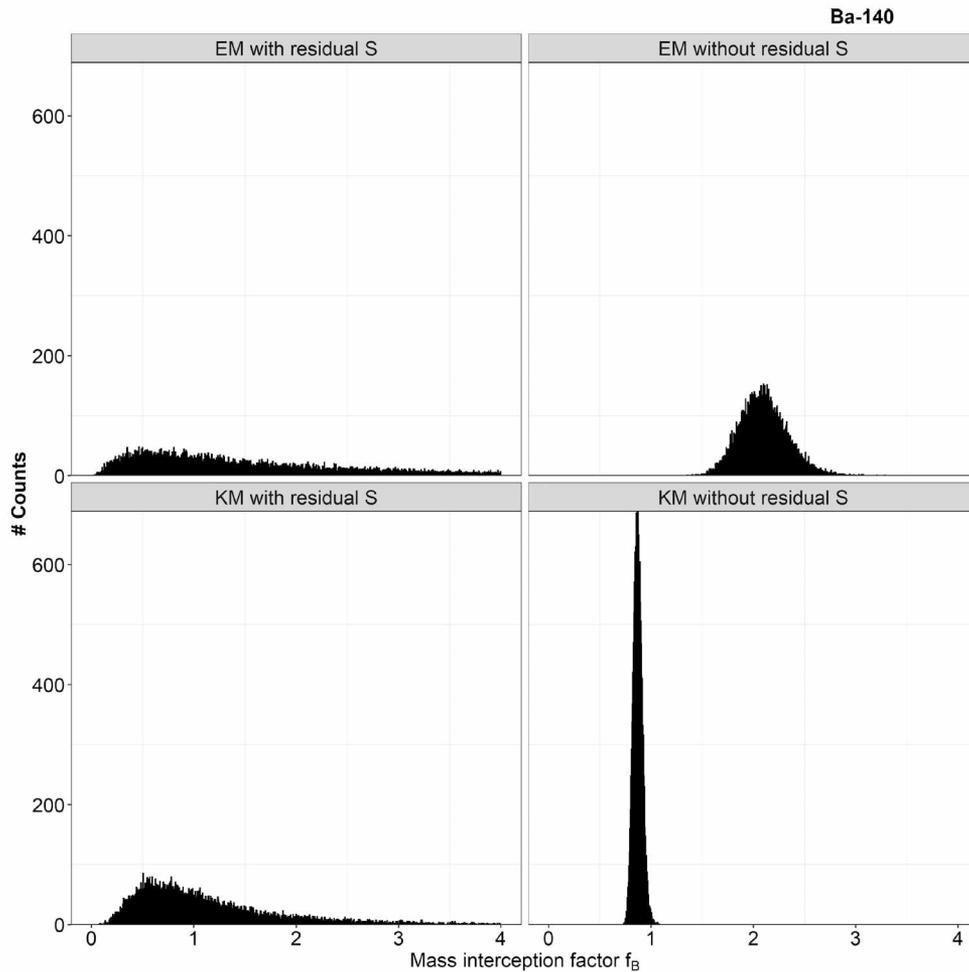


Fig 2. Example of posterior distributions of f_B obtained with and without propagating S for EM (top) and KM (bottom).

is shown in Fig. 4, in which CV_{95} from pdfs obtained with (right) and without propagating S (left) are plotted. For Cs-137 and KM, the dispersion of the model output around the median for Cs-137 without S is rather large compared to the other elements. From Fig. 5, it can be observed, as expected, that the posterior distributions obtained with single S have smaller dispersion around the median value compared to these obtained with one unique S . This is due to the variability arising from the chemical valence of the pollutant. The dispersion of the f_B values around the median increases by a factor 4 for both KM and for EM.

4. DISCUSSION

A methodology based on the analysis of residuals within a Bayesian framework has been applied to quantify conceptual model uncertainty of two process-based models, EM and KM, for modeling the process of interception of wet deposited pollutants. In particular, this methodology discriminates between conceptual model uncertainty and parameter uncertainty. In this way, the impact of conceptual model uncertainty is explicitly considered and not implicitly attributed to parameter uncertainty. The applied methodology combines the use of

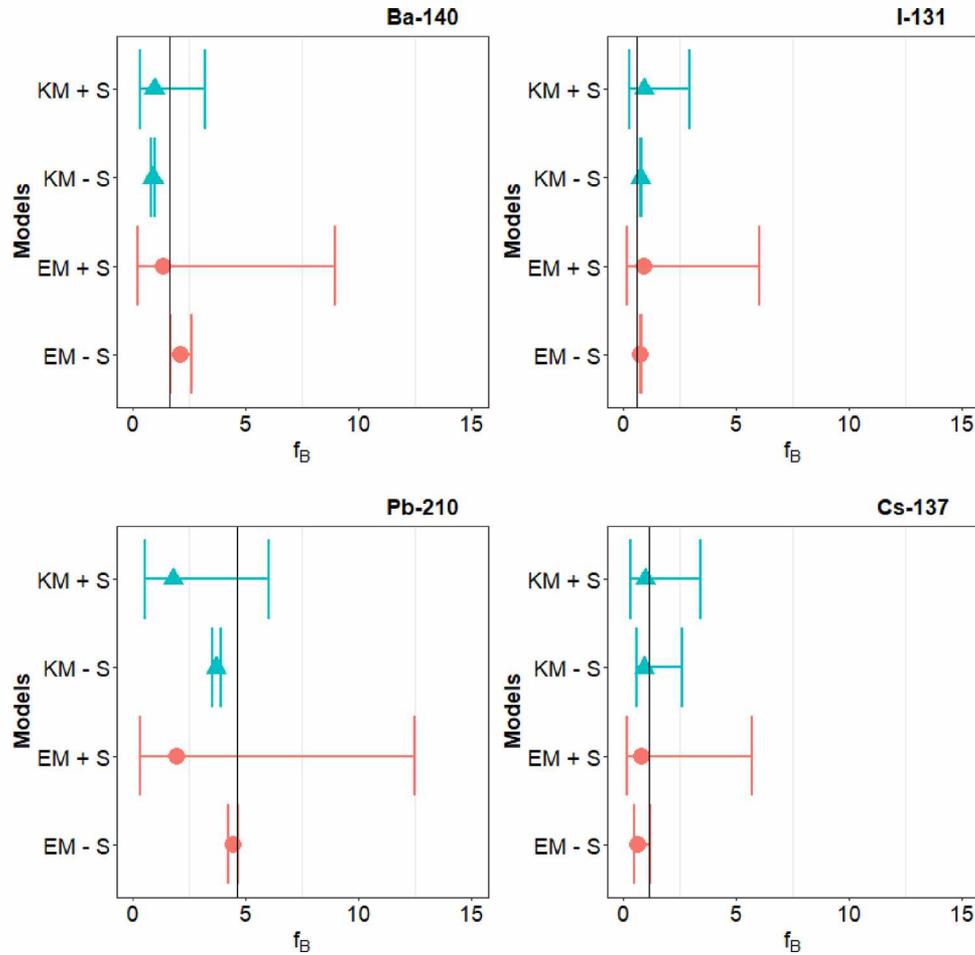


Fig 3. 95% credibility intervals for EM (circle) and KM (triangle) obtained with and without propagating the conceptual model uncertainty, that is, (EM +S, KM +S) and (EM -S, KM -S), respectively. Black vertical line indicates the experimentally obtained f_B .

residual analysis and the Bayesian approach, which allows for quantification of conceptual model uncertainty efficiently and includes expert judgement in a formalized way. The coefficient of variation CV_{95} is the metric applied to quantify impact of conceptual model uncertainty and it includes 95th percentile in its estimation. The 95th percentile is important in radioecological assessments, especially for regulatory purposes (ICRP, 2006). This metric proves to be adequate not only to quantify how conceptual model uncertainty affects the outputs of EM and KM, respectively, but also to quantify how natural variability contributes to modeling outputs.

Nevertheless, there are fundamental assumptions that underlie the study, namely that the Bayesian residual is a suitable measure of total uncertainty of EM and KM and that the data used are representative for the scientific question under study.

The Bayesian methodology allows for straightforwardly propagating the uncertainties related to the parameters and model structure to the model output. This has the advantage that typical assumptions on additivity and correlations between the different uncertainty contributions need not to be carried out.

If no adequate measurement and sampling information is available, the uncertainty of input variables that depend on the experimental set-up (such as H , I , B in the present study) cannot be properly quantified.

Dedicated effort already during the experimental procedure to gather information about uncertainty related to measurements and sampling is necessary. However, such information is not always gathered because data are often collected for specific purposes and are not targeted to modeling activities. In several studies—from various

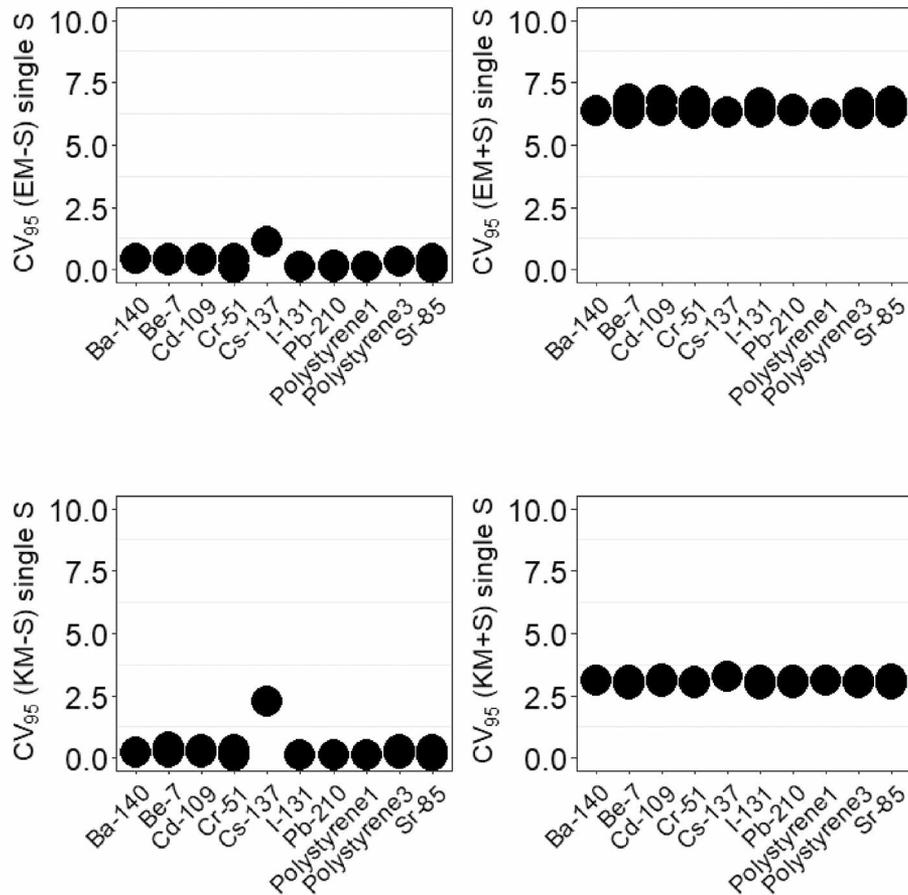


Fig 4. CV_{95} obtained for EM (top) and KM (bottom) by considering pdf of f_B with propagated S (right) and the CV_{95} obtained for EM and KM by considering pdf of f_B without propagated S (left).

environmental disciplines —the same issue is reported for quantities like evapotranspiration, temperature, soil water content, and so on (Radwan et al., 2004) (Højberg & Refsgaard, 2005) (Jin et al., 2010) (Alderman & Stanfill, 2017). Nevertheless, methods to quantify uncertainties related to sampling activities and/or measurements exist (Kirkup & Frenkel, 2006). In particular, assessing the performance of measurement sensors (Chen, Chen, & Chen, 2018) (Damgaard, 2020) is crucial to determine the uncertainty due to measurements. In studies where input uncertainty is considered and discussed at least qualitatively, more exhaustiveness in the determination of the total uncertainty budget of a model is obtained (see Section 3.2 in Gondwe, Merediz-Alonso, & Bauer-Gottwein, 2011; Radwan et al., 2004).

In environmental models, data available not only should provide as much information as possible re-

garding input quantities but also need to be representative of the situation being analyzed that is cover the entire validity domain of the model. Instead, data may be rather limited: for the present radioecological study the 440 observations available with large range of variation of quantities such as I or H and for different types of radionuclides are more the exception rather than the rule.

For Bayesian analysis the choice of priors requires information on the possible range of variation of the quantities involved and the choice of most sensitive parameters needs to be outweighed based also on the possibility to build a probability distribution function. A more formal sensitivity analysis on all input quantities (Saltelli, 2002) is an interesting perspective when the aim to characterize the total uncertainty budget.

In the field of radioecology not so many different conceptual models for a given transfer process may

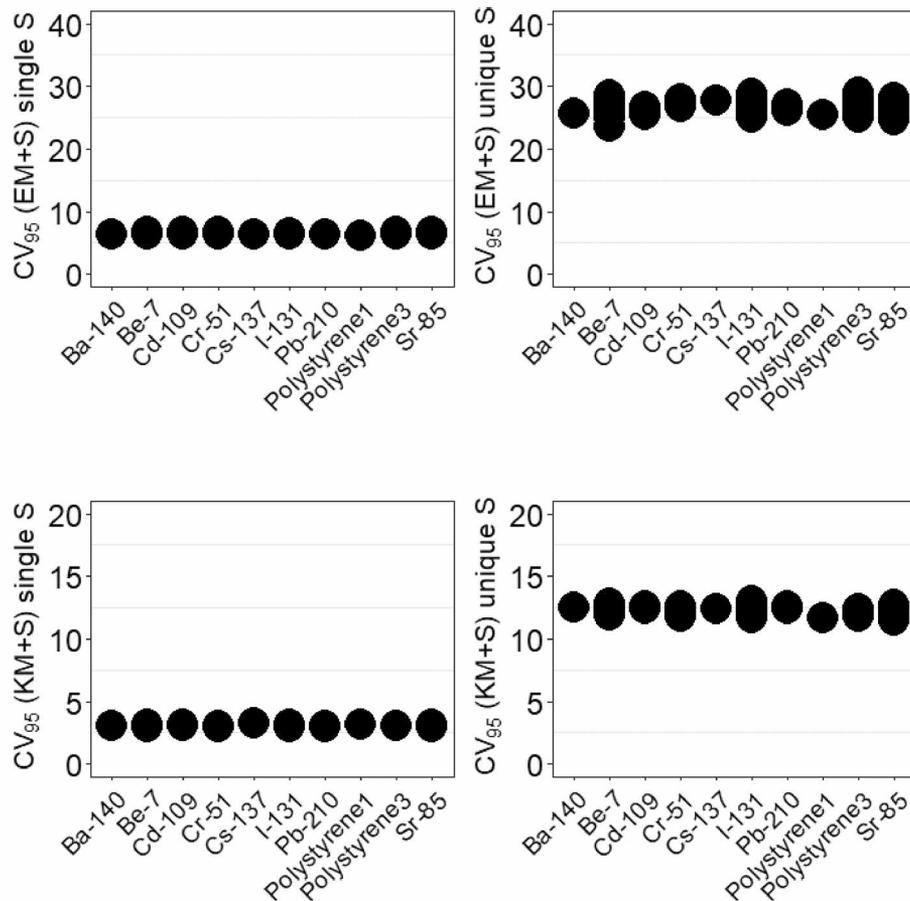


Fig 5. CV_{95} obtained for EM (top) and KM (bottom) by considering posterior distributions of f_B with propagated single S (left) and propagated unique S (right).

be considered or are available (Urso et al., 2015) as can be the case instead in other disciplines (e.g., Gondwe et al., 2011; Højberg & Refsgaard, 2005). Indeed, the availability of two different process-based models for describing the process of interception of wet deposited pollutants on herbaceous species is rather unique and motivated the present study. The results from various studies (Gondwe et al., 2011; Jin et al., 2010; Refsgaard et al., 2006) indicate that conceptual model uncertainty contributes even more than parameter uncertainty to the total uncertainty budget of a model in agreement with the present findings. Also, the analysis of conceptual model uncertainty improves model understanding and confidence in model predictability. It is especially when data are not available and hence models are needed for prognostic assessment that improved understanding of the uncertainty contributions is necessary. The present case is very specific and more

modeling examples and research is required before generalizations valid for radioecological models can be made. However, the developed methodology can be certainly applied to other radioecological models.

5. CONCLUSION

The main finding is that KM describes experimental data better than EM in agreement with (Gonze & Sy, 2016). However, in (Gonze & Sy, 2016) the predictive model performance is carried out by means of posterior predictive loss criterion (PPLC). The present study suggests model performance testing by looking at how the model structure affects the gap between the observations and the corresponding predictions. In fact, the substantial differences in the EM and KM mathematical structures and the fact that the same set of data was used for both models

justifies the consideration of conceptual model uncertainty as being the dominant contribution to residuals. This analysis shows that for the majority of data points the process-based approach, which models the interception process as a kinetic process, provides reduced structural uncertainty (and hence better agreement) than the modeling of the interception process under equilibrium conditions.

The main conclusions are:

- the 95% credibility intervals are larger for EM than KM for the majority of the datapoints.
- CV_{95} values obtained with and without single S show that, on average, conceptual model uncertainty is larger for EM than KM by at least a factor 10.
- CV_{95} values obtained using single S and unique S show that the effect of variability related to chemical valence on EM and KM output affects the dispersion of the model outputs around the mean value by a factor 4.

ACKNOWLEDGMENTS

This work has been supported by the TERRITORIES project, part of the CONCERT European Joint Programme funded from the Euratom research and training programme 2014–2018 under grant agreement No. 662287. The present publication reflects only the authors' view. Responsibility for the information and views expressed therein lies entirely with the authors. The European Commission is not responsible for any use that may be made of the information it contains. Open Access funding enabled and organized by Projekt DEAL.

REFERENCES

- Alderman, P. D., & Stanfill, B. (2017). Quantifying model-structure-and parameter-driven uncertainties in spring wheat phenology prediction with Bayesian analysis. *European Journal of Agronomy*, 88, 1–9.
- Angeletti, L., & Levi, E. (1977a). *A study of the wet deposit and foliar uptake of iodine and strontium on rye-grass and clover*. CEA-R-4897. Paris, France: CEA Centre d'Etudes Nucleaires de Fontenay-aux-Roses.
- Angeletti, L., & Levi, E. (1977b). *A comparative study of transfer factors of water, iodine and strontium on rye-grass and clover. Development of a model of evaluation of the limits of foliar contamination by wet deposit*. CEA-R-4860. Paris, France: CEA Centre d'Etudes Nucleaires de Fontenay-aux-Roses.
- Ascough II, J. C., Maier, H. R., Ravalico, J. K., & Strudley, M. W. (2008). Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecological Modelling*, 219(3–4), 383–399.
- Bengtsson, S. B., Gårdenäs, A. I., Eriksson, J., Vinichuk, M., & Rosén, K. (2014). Interception and retention of wet-deposited radiocaesium and radiostrontium on a ley mixture of grass and clover. *Science of the Total Environment*, 497, 412–419.
- Box, G. E. P., & Tiao, G. C. (2011). *Bayesian inference in statistical analysis* (Vol. 40). New York: John Wiley & Sons.
- Chadwick, R. C., & Chamberlain, A. C. (1970). Field loss of radionuclides from grass. *Atmospheric Environment* (1967), 4(1), 51–56.
- Chen, L.-H., Chen, J., & Chen, C. (2018). Effect of environmental measurement uncertainty on prediction of evapotranspiration. *Atmosphere*, 9(10), 400.
- Clark, L. (1940). A new species of Nardia. *The Bryologist*, 43(2), 29–32.
- Damgaard, C. (2020). Measurement uncertainty in ecological and environmental models. *Trends in Ecology & Evolution*, 35(10), 871–873.
- Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 45–70.
- Engeland, K., Xu, C.-Y., & Gottschalk, L. (2005). Assessing uncertainties in a conceptual water balance model using Bayesian methodology/Estimation bayésienne des incertitudes au sein d'une modélisation conceptuelle de bilan hydrologique. *Hydrological Sciences Journal*, 50(1). <https://doi.org/10.1623/hysj.50.1.45.56334>
- EPA. (2009). *Guidance on the development, evaluation, and application of environmental models*. EPA/100/K-09/003. Washington, DC: Environmental Protection Agency.
- Gondwe, B., Merediz-Alonso, G., & Bauer-Gottwein, P. (2011). The influence of conceptual model uncertainty on management decisions for a groundwater-dependent ecosystem in karst. *Journal of Hydrology*, 400(1–2), 24–40.
- Gonze, M.-A., & Sy, M. M. (2016). Interception of wet deposited atmospheric pollutants by herbaceous vegetation: Data review and modelling. *Science of the Total Environment*, 565, 49–67.
- Grunsky, E. (2002). R: a data analysis and statistical programming environment—An emerging tool for the geosciences. *Computers & Geosciences*, 28(10), 1219–1222.
- Hoffman, F. O., & Hammonds, J. S. (1994). Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk analysis*, 14(5), 707–712. doi:<https://doi.org/10.1111/j.1539-6924.1994.tb00281.x>
- Hoffman, F. O., Thiessen, K. M., Frank, M. L., & Blaylock, B. G. (1992). Quantification of the interception and initial retention of radioactive contaminants deposited on pasture grass by simulated rain. *Atmospheric Environment. Part A. General Topics*, 26(18), 3313–3321.
- Hoffman, F. O., Thiessen, K. M., & Rael, R. M. (1995). Comparison of interception and initial retention of wet-deposited contaminants on leaves of different vegetation types. *Atmospheric Environment*, 29(15), 1771–1775.
- Højberg, A., & Refsgaard, J. (2005). Model uncertainty—parameter uncertainty versus conceptual models. *Water Science and Technology*, 52(6), 177–186.
- Horton, R. E. (1919). Rainfall interception. *Monthly Weather Review*, 47(9), 603–623.
- IAEA. (2011). *Radioactive particles in the environment: sources, particle characteristics, and analytical techniques*. IAEA-TECDOC No. 1663. Vienna, Austria: International Atomic Energy Agency.
- ICRP. (2006). Assessing dose of the representative person for the purpose of the radiation protection of the public. ICRP Publication 101a. *Annals of the ICRP* 36, 3.
- Jacob, P., Müller, H., Pröhl, G., Voigt, G., Berg, D., Paretzke, H. G., & Regulla, D. (1993). Environmental behaviour of radionuclides deposited after the reactor accident of Chernobyl and related exposures. *Radiation and environmental biophysics*, 32(3), 193–207.

- Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106(4), 620.
- Jin, X., Xu, C.-Y., Zhang, Q., & Singh, V. P. (2010). Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *Journal of Hydrology*, 383(3-4), 147–155.
- Jouven, M., Carrère, P., & Baumont, R. (2006). Model predicting dynamics of biomass, structure and digestibility of herbage in managed permanent pastures. 1. Model description. *Grass and Forage Science*, 61(2), 112–124.
- Kinnersley, R. P., Goddard, A. J. H., Minski, M. J., & Shaw, G. (1997). Interception of caesium-contaminated rain by vegetation. *Atmospheric Environment*, 31(8), 1137–1145.
- Kirchner, G., & Steiner, M. (2008). Uncertainties in radioecological assessment models—Their nature and approaches to reduce them. *Applied Radiation and Isotopes*, 66(11), 1750–1753. <https://doi.org/10.1016/j.apradiso.2007.09.021>
- Kirkup, L., & Frenkel, R. B. (2006). *An introduction to uncertainty in measurement: using the GUM (guide to the expression of uncertainty in measurement..* Cambridge, UK: Cambridge University Press.
- Leelössy, Á., Lagzi, I., Kovács, A., & Mészáros, R. (2018). A review of numerical models to predict the atmospheric dispersion of radionuclides. *Journal of Environmental Radioactivity*, 182, 20–33. <https://doi.org/10.1016/j.jenvrad.2017.11.009>
- Lindenschmidt, K.-E., Fleischbein, K., & Baborowski, M. (2007). Structural uncertainty in a river water quality modelling system. *Ecological Modelling*, 204(3-4), 289–300.
- McAllister, M., & Kirchner, C. (2002). Accounting for structural uncertainty to facilitate precautionary fishery management: illustration with Namibian orange roughy. *Bulletin of Marine Science*, 70(2), 499–540.
- Merriam, R. A. (1961). Surface water storage on annual ryegrass. *Journal of Geophysical Research*, 66(6), 1833–1838.
- Neuman, S. P., Wierenga, P. J., & Nicholson, T. J. (2003). A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites: Division of Systems Analysis and Regulatory Effectiveness, Office of Nuclear Regulatory Research. NUREG/CR-6805. Washington, DC: U.S. Nuclear Regulatory Commission.
- Özkaynak, H., Frey, H. C., Burke, J., & Pinder, R. W. (2009). Analysis of coupled model uncertainties in source-to-dose modeling of human exposures to ambient air pollution: A PM2.5 case study. *Atmospheric Environment*, 43(9), 1641–1649.
- Petersen, A. C., Janssen, P. H. M., & Van der Sluijs, J. P. (2003). *RIVM/MNP Guidance for uncertainty assessment and communication*. Utrecht, The Netherlands: Utrecht University.
- Pröhl, G. (2003). Chapter 4 Radioactivity in the terrestrial environment. In E. M. Scott (Ed.), *Modelling radioactivity in the environment* (Vol. 4, pp. 87–108). Amsterdam, The Netherlands: Elsevier.
- Radwan, M., Willems, P., & Berlamont, J. (2004). Sensitivity and uncertainty analysis for river quality modelling. *Journal of Hydroinformatics*, 6(2), 83–99.
- Refsgaard, J. C., Van der Sluijs, J. P., Brown, J., & Van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, 29(11), 1586–1597.
- Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process—a framework and guidance. *Environmental Modelling & Software*, 22(11), 1543–1556.
- Rowe, W. D. (1994). Understanding Uncertainty. *Risk Analysis*, 14(5), 743–750. <https://doi.org/10.1111/j.1539-6924.1994.tb00284.x>
- Salbu, B. (2016). Environmental impact and risk assessments and key factors contributing to the overall uncertainties. *Journal of Environmental Radioactivity*, 151, 352–360.
- Saltelli, A. (2002). Sensitivity analysis for importance assessment. *Risk Analysis*, 22(3), 579–590.
- Simon-Cornu, M., Beaugelin-Seiller, K., Boyer, P., Calmon, P., Garcia-Sanchez, L., Murlon, C., Nicoulaud, V., Sy, M. M., & Gonze, M. A. (2015). Evaluating variability and uncertainty in radiological impact assessment using SYMBIOSE. *Journal of Environmental Radioactivity*, 139, 91–102.
- Symonds, M. R. E., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, 65(1), 13–21.
- Urso, L., Hartmann, P., Diener, A., Steiner, M., & Vives i Batlle, J. (2015). *Deliverable 3.4 Report on the feasibility of improving radioecological models using a process-oriented approach*. STAR project, European Commission Contract Fission-2010-3.5.1-269672 Strategy for Allied Radioecology.
- Walker, W. E., Harremoës, P., Rotmans, J., Van Der Sluijs, J. P., Van Asselt, M. B. A., Janssen, P., & Krayen von Krauss, M. P. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, 4(1), 5–17.
- Wallach, D., Nissanka, S. P., Karunaratne, A. S., Weerakoon, W. M. W., Thorburn, P. J., Boote, K. J., & Jones, J. W. (2017). Accounting for both parameter and model structure uncertainty in crop model predictions of phenology: A case study on rice. *European Journal of Agronomy*, 88, 53–62.
- Walsh, L., & Kaiser, J. C. (2011). Multi-model inference of adult and childhood leukaemia excess relative risks based on the Japanese A-bomb survivors mortality data (1950–2000). *Radiation and Environmental Biophysics*, 50(1), 21–35.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Excel Table with set of input data used for calculations. Excel Table with results of analysis with single S (i.e. S_i) and analysis with unique S .

Supporting Information
Supporting Information