

This project has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 662287.



EJP-CONCERT

European Joint Programme for the Integration of Radiation Protection Research

H2020 – 662287

D9.61 – Guidance to select level of complexity

Lead Author: Justin Brown (DSA), Alla Dvorzhak (CIEMAT)

With contributions from: Juan C. Mora, Danyl Pérez-Sanchez (CIEMAT), Marko Kaasik, Alan Tkaczyk (UT), Ali Hosseini, Mikhail Iosjpe, Jelena Popic (DSA), Justin Smith (PHE), Jordi Vives i Batlle, Talal Almahayni, Nathalie Vanhoudt (SCK•CEN), Marc-André Gonze, Philippe Calmon, Laureline Février (IRSN), Philipp Hartmann, Martin Steiner, Laura Urso (BfS), Deborah Oughton, Ole Christian Lind, Brit Salbu (NMBU)

Reviewers: Juan Carlos Mora (CIEMAT), Marie Simon-Cornu (IRSN), and CONCERT coordination team

Work package / Task	WP 9	T9.3 (TERRITORIES)	SST 9.3.1.2
Deliverable nature:	Report		
Dissemination level: (Confidentiality)	Public		
Contractual delivery date:	M50		
Actual delivery date:	M50		
Version:	1.0		
Total number of pages:	171		
Keywords:	Radioecological models (complexity, improvement and performance); Fukushima; NORM; Environmental impact		
Approved by the coordinator:	M50		
Submitted to EC by the coordinator:	M50		

Disclaimer:

The information and views set out in this report are those of the author(s). The European Commission may not be held responsible for the use that may be made of the information contained therein.

Abstract

Models in radioecology, as in other fields, have several purposes, the most important being the prediction of the behavior of radionuclides in different ecosystems and the understanding of the processes driving that behavior. These models are often finally used for regulatory purposes by transforming the values to a limiting quantity, such as effective dose or absorbed dose to demonstrate the protection of humans or biota respectively. To account for the consideration that several processes are not perfectly known, a considerable overestimation of the predictions is normally included in the models. Moreover, the final estimations of doses are directly proportional to the estimations of activity concentrations in the environment. Obviously, for many applications, only models can be used for prognosis such as predicting future activity concentrations. For that reason, and to avoid undue restrictions caused by poor results of the models, improvement of models is desirable and a continuous effort in this direction is needed. In this report, an example of a methodology which can be used to systematically improve the models is presented by providing a conceptual overview of the system through the use of Interaction Matrices and Features, Events and Processes.

For the developers and the end users of the models, objective indicators to show whether a model is improved or not are desirable. Adapted from other fields, notably from the meteorological sciences, a methodology combining quantitative and qualitative indicators is elaborated. These indicators are used together with measured data in the different ecosystems where the comparison is needed.

Finally, in those sites included in the Territories Library Database where a compilation of measured data was included, a comparison of widely used models (usually simpler) with more advanced models (usually more complex) has been implemented. Specifically several models have been applied in the Norwegian Fen site (NORM), in the Belgian NORM site, in the Fukushima forests contaminated by the FDNPP 2011 accident and in the West Cumbrian beaches, contaminated by releases from the Sellafield reprocessing facility. In all the cases, a discussion on the models, together with a comparison of the indicators applied to each model used at every site, was included.

Altogether, this report can be regarded as a methodology to improve and show objectively the improvement of models applied to real case studies of long-term situations where contamination exists (often referred to as legacy sites). Applications in different situations can be seen as examples of implementing this process.

TABLE OF CONTENTS

1	Executive summary	7
2	Introduction.....	8
2.1	Background and aims	8
2.2	Approach and report contents	9
2.3	Expectation and purposes	11
3	HOW TO IMPROVE A MODEL	13
3.1	Interaction matrices – FEPs analysis	13
3.1.1	Principles	13
3.1.2	Example: Interaction matrix for a forest ecosystem	15
3.2	Need for a methodology to quantify model improvement in radioecology.....	18
3.3	Model performance	19
3.3.1	What does model performance mean?	19
3.3.2	How to evaluate model performance	20
3.4	Overview of numerical tools for evaluating model performance in environmental sciences..	21
3.4.1	Direct value comparison.....	22
3.4.2	Coupling real and modelled values	22
3.4.3	Dealing with time series	24
3.5	Suggested metrics for radioecology.....	25
3.6	Availability of experimental data	26
3.6.1	Methods for data splitting.....	27
3.6.2	Model performance and model uncertainties	27
3.6.3	Proposed methodology in radioecology	28
3.6.4	Description of suggested qualitative criteria	28
3.6.5	Description of suggested numerical measure.....	29
3.6.6	Performance table for combining qualitative and quantitative indicators.....	29
4	DESCRIPTION OF THE MODELS.....	31
4.1	EXISTING/WIDELY APPLIED MODELS.....	32
4.1.1	CROM.....	32
4.1.2	NORMALYSA	34
4.1.3	FORESTCROM	38
4.1.4	TREE4-simple	44
4.1.5	NRPA box modelling approach (ARCTICMAR)	46
4.2	MORE ADVANCED MODELS.....	50
4.2.1	GRANIS.....	50
4.2.2	ECOFOR.....	52
4.2.3	TALAL	54
4.2.4	TREE4-advanced	57
4.2.5	AMIS	61
5	Case study: NORWEGIAN FEN FOREST SITE	64
5.1	Introduction to the case study	64
5.2	Application of the models to case study	65
5.2.1	CROM.....	65
5.2.2	NORMALYSA	67
5.2.3	GRANIS.....	73

6	Case study: BELGIAN NORM SITE	76
6.1	Introduction to the case study	76
6.2	Application of the models to case study	77
6.2.1	FORESTCROM	77
6.2.2	ECOFOR.....	81
6.2.3	TALAL	88
7	Case study: FUKUSHIMA.....	93
7.1	Introduction to the case study	93
7.2	Application of the models to case study	94
7.2.1	FORESTCROM	94
7.2.2	TREE4-simple	99
7.2.3	TREE4-advanced	102
8	Case study: SELLAFIELD	106
8.1	Introduction to the case study	106
8.2	Application of the models to case study	107
8.2.1	ARCTICMAR and AMIS models. Quantifying improvement the local model (AMIS) versus the NRPA box model (ARTICMAR)	107
8.2.2	AMIS model to quantify the fate of “hot particles” after long-term exposure in the Sellafield intertidal beach region.....	112
9	Discussion	120
9.1	Good modelling practice	121
9.1.1	A tiered approach to assessments	121
9.1.2	The steps in model development and implementation	122
9.1.3	Quality Assurance.....	127
9.2	Performance assessment analysis of the models. Comparison of the models.....	128
10	Recommendations.....	133
10.1	Selection of a model (prior to performing assessment).....	133
10.2	Use of a model in an assessment	134
10.3	Developing a model.....	134
11	Bibliography.....	136
12	Appendices	144
12.1	Performance assessment analysis. Old vs New model comparison.	144
12.1.1	NORWEGIAN FEN FOREST SITE.....	144
12.1.2	BELGIAN NORM SITE	148
12.1.3	FUKUSHIMA.....	152
12.1.4	SELLAFIELD.....	156
12.2	Uncertainty assessment (Uncertainty matrix, Model Pedigree).....	159
12.2.1	The dimensions of uncertainty	159
12.2.2	Approaches available for uncertainty assessment.....	160
12.2.3	Selection of appropriate methodology	168
12.2.4	References for Appendix 12.2	169

Abbreviations

<i>AL</i>	<i>Activity level</i>
<i>CAC</i>	<i>Cation Exchange Capacity</i>
<i>CCF</i>	<i>Cross-correlation function</i>
<i>CR</i>	<i>Concentration ratio</i>
<i>BIAS</i>	<i>Metric bias (Mean of the residuals)</i>
<i>FEPs</i>	<i>Features-Event Processes</i>
<i>FDNPP</i>	<i>Fukushima-Daiichi Nuclear Power Plant</i>
<i>LHS</i>	<i>Latin hypercube sampling</i>
<i>LAI</i>	<i>Leaf area index</i>
<i>MAE</i>	<i>Mean absolute error</i>
<i>MG</i>	<i>Geometric mean bias</i>
<i>MSE</i>	<i>Mean squared error</i>
<i>MSLE</i>	<i>Logarithmic Mean Square Error</i>
<i>NORM</i>	<i>Naturally occurring radioactive material</i>
<i>RMSE</i>	<i>Root mean squared error</i>
<i>RMSLE</i>	<i>Root mean squared logarithmic error</i>
<i>RN</i>	<i>Radionuclide</i>
<i>PPMC</i>	<i>Pearson product moment correlation coefficient</i>
<i>QA</i>	<i>Quality assurance</i>
<i>QQ</i>	<i>Quantile-quantile</i>
<i>UA</i>	<i>Uncertainty Analysis</i>

1 EXECUTIVE SUMMARY

Contamination of the environment with radioactive material can lead to the exposure of both humans and non-human biota. To protect them from the damage that this exposure to radioactivity can cause, the effective dose or absorbed dose needs to be quantified in a process of a dose assessment. These quantities are not directly measured, but, on the contrary, are based on measurements of many other quantities, such as the external exposure quantified by measuring dose equivalent, or the activity concentrations in the different compartments of the environment. For this reason, to quantify the level of any potential exposure, the characteristics of the contamination must be identified, in terms of location or spatial extent, temporal variation, quantity, and radionuclides involved. Although doses in retrospective assessments should be based on measured concentrations, often those measurements need to be complemented with model values, due to the impossibility of performing measurements, because the empirical data do not cover the total extent of the contamination, or just because measured levels are below detection limits. Prospective assessments of future exposures, however, are necessary in many applications, as safety studies prior to commissioning a new installation or to select appropriate techniques to remediate a contaminated zone. In these situations, measurements are not practicable, and the results must be based on models. In such cases, models are necessary to predict the dispersion of radionuclides through the environment and to calculate activity concentrations or external exposures in the relevant environmental media. In order to verify the validity of the predictions, the employed models should be compared against real measurements in a range as wide as possible.

Choosing an appropriate, or 'fit-for-purpose', model is not necessarily straightforward. It depends on the assessment context and stakeholder requirements. For example, for fundamental research into environmental transfer, a detailed as possible process-based model is required to advance the state of the art in knowledge is required. On the other hand, for regulatory control a simpler but cautious model may be most suitable; such a model could be based on concentration ratios between environmental compartments derived from experimental studies. In all the cases, the level of uncertainty associated with the model outputs should be addressed and be reduced to a level that is fit-for-purpose.

This report comprises a technical guidance including examples and methods to achieve more advanced models, as well as recommendations on a methodological approach to identify models that are fit-for-purpose in terms of complexity and uncertainty. Additionally a methodology is provided to measure quantitatively and qualitatively the results of the different models. It is recommended that a qualitative analysis of model performance should include consideration of the number of model parameters, number of separate processes, and the application of the model to empirical datasets. For a quantitative analysis, a key component is to compare the correspondence between modelled and empirical datasets; recommendations for achieving this correspondence are provided for radioecological models. It is proposed that the primary statistical indicator to use is the root mean squared error (or root mean squared logarithmic error for sparse data), although other indicators are described, such as bias, geometric mean bias and, for time series, the cross-correlation function. It is suggested that these various measures of model performance can usefully be included in a performance table that lists the measures, what they tell us about the model and their limitations.

Practical examples of the application of this methodology are provided for test cases in which the outputs of existing, widely applied models are compared with empirical data, which have been compiled within the TERRITORIES Library Database (the TLD) described elsewhere (Smith *et al.*, 2019) for several radioactively contaminated sites. This comparison is then repeated for more advanced, process-based models. A comparison of the performance of the existing and advanced models is made

and reasons for model improvement, or indeed deterioration, are explained. Examples of completed performance tables are given for each test case.

The report also considers methods for developing models from scratch or improving existing model performance through the application of good modelling practice that includes the following steps (1) Model Study Plan, (2) Data and conceptualisation, (3) Model set-up, (4) Calibration and validation and (5) Simulation and evaluation. These steps include valuable advice such as clearly defining, at the problem formulation stage, the criteria against which the model will be judged, the use of Features-Event Processes analysis to identify the main transfer processes, defining model scope and characterising uncertainties in model outputs. Uncertainty assessment will ideally constitute an important intrinsic component in establishing whether a model is fit-for-purpose and the appropriate level of complexity requirements for a given model. Approaches and tools used for uncertainty analysis are dealt with in detail elsewhere (Urso et al, 2019).

The quality assurance (QA) procedures that have been followed during model development will clearly impact on whether the outputs of a model are robust and reliable and hence if it is fit-for-purpose.

The lessons learnt that have emerged from the process outlined above, along with more general considerations, arising from the experience of the report authors in relation to risk assessment, have been assimilated. This has allowed the development of guidance and recommendations within the context of establishing a germane fit-for-purpose level of complexity for model development. A key recommendation of this document is to select a model that is adequate to the objective. This includes selecting a model appropriately conservative in those cases where the protection of the humans or biota is regarded. This is often the objective when looking for an optimum option in remediating a contaminated site. However, an overly conservative approach will unduly produce an over-cost associated with the evaluations. For this reason, more realistic models (fitted to real measurements for a maximally wide range of possible situations) are preferred, given that every parameter needed for that model has been measured in advance. This point restricts many times the application of very complex models, which need an elevated number of parameters that are not usually measured. However, in those cases where a high precision is desirable, as can be the case in research applications where a full understanding of the underlying processes is being pursued, conservatism is not an advantage, and all the resources and efforts needed to characterize every parameter included in the model should be provided in order to achieve more accurate results. Finally, the methodology used to establish the performance of different models is a tool that can help in the decision regarding the selection of the desired model for a given application.

2 INTRODUCTION

2.1 Background and aims

Quantifying the potential impact of exposure to radiation on humans and the environment in a reliable and robust manner is a central requirement for practitioners of radiological protection. This has typically been addressed by linking sources of radioactivity to exposure and then to (biological) effect/impact within a structured framework as evidenced by the numerous tools and methodologies available through international efforts (ICRP, 2007; IAEA, 2001). When feasible, it is ideally based on information directly obtained from radiological characterization of contaminated territories. Nevertheless, to answer some questions, e.g. for predicting evolution, or to anticipate potential effect of not-yet implemented remediation action, such information gained from monitoring is not sufficient and mathematical models can be used to simulate the behaviour and fate of radionuclides within a

given environmental system. These might be defined in their simplest form as single parameter equations, for example, characterising the ratio between the radionuclide activity concentrations in environmental compartments and, in more complex forms, such as those constituting multi-parameter models, requiring solutions to nonlinear partial differential equations. An elementary example of the latter would involve simulating the migration of solutes in unsaturated soils with solid matter interaction that typically, for realistic soil wetting cycles and seasonally changing evapotranspiration, might require numerical integration.

The selection of the most appropriate model is not necessarily a straightforward proposition. It seems self-evident that the choice would be, among other considerations, (assessment) context-dependent. A research group interested in gaining a mechanistic understanding of the processes involved in the biological uptake of a given radioelement would clearly have quite different criteria to evaluate what model is 'fit-for-purpose' compared to a regulatory authority, which might simply be concerned with modelling biological transfer of the same radioelement, in a reliable and defensible manner, as a minor part of a comprehensive discharge authorisation. It also seems self-evident that the assessment context should also be informed by the requirements of the stakeholders involved in the process. The modern practice of including stakeholders at every stage of an assessment are well established and documented (Refsgaard *et al.*, 2007) rendering further justification for this practice, in the present report, redundant. Nonetheless, many challenges remain when addressing the practicalities of this subject and bespoke guidance on model selection within the field of radiological protection and radioecological assessment has not received serious attention previously.

One of the overarching aims of the TERRITORIES project was to build a methodological approach to reduce uncertainties to a fit-for-purpose level in support of the decision making process. This could be achieved by encompassing stakeholder involvement to assess and manage, in a graded way, (effective and absorbed) doses associated with existing¹/long-lasting exposure situations of both human and wildlife populations. The process was envisaged to provide fit-for-purpose information for environmental diagnosis and prognosis of long-lasting exposure situations, and in relation to the environmental behaviour of the radionuclides, taking account of stakeholder's concerns. A key component of this effort was anticipated to involve the application of models at sites of interest, primarily selected to cover various existing exposure situations, to provide simulations of the behaviour and fate of radionuclides and resultant impacts on the environment. A technical guidance with recommendations about the desirable fit-for-purpose level of complexity for model development was considered important in attaining this goal. **The aim of this report is to address this theme through a systematic appraisal of the types of information a modeller, i.e. a person or an organisation that works with or develops a model, would typically have to hand and the considerations that are made during model application and development through the employment of case studies.**

2.2 Approach and report contents

The approach adopted in this report is admittedly mainly heuristic, being framed by the information that was available to the TERRITORIES project and by the models that have been typically used or are under development by the groups involved in this study.

¹ Existing exposure situation, Defined in ICRP (2007) as 'exposure situations that already exist when a decision on control has to be taken, such as those caused by natural background radiation.'

The point of departure has been the empirical dataset collated in the TERRITORIES Library Database (TLD) as presented elsewhere (Smith *et al.* 2019). To assist the reader, a brief overview of the empirical datasets available for the cases selected for further consideration through modelling are provided in relevant chapters of the report (i.e. Sections 5 to 8). The cases include both terrestrial forest systems (contaminated by accidental releases for sites in Fukushima and by NORM² for sites in Fen-Norway and Belgium) and marine systems (contaminated by routine releases for a Sellafield coastal site in the UK).

As noted by Urso *et al.* (2019), there are common determination endpoints, i.e. model outputs, associated with radiological risk characterisation and the employment of models simulating the behaviour and fate of radionuclides. Typical determinations are in the form of radionuclide contamination densities (Bq m⁻²), activity concentrations in environmental media (Bq L⁻¹ or Bq kg⁻¹) and some form of dose determination such as absorbed dose-rate (Gy h⁻¹). Because these model outputs are compatible with quantities measurable in the field, as evidenced by the data collated as such in the TLD, they have been selected to delineate the scope of the analysis conducted. In other words, the scope of the subsequent discussions is bounded by radionuclide concentration determinations and (empirically-testable) dose (rate) determinations but does not encroach upon the meaning of these in terms of detriment to human health or environmental impact.

It was, furthermore, considered useful to address the broader context of environmental risk assessment by posing the question: “what are the typical purposes and expectations associated with the implementation of models?” The plan was to achieve this by eliciting the stakeholders perspective through dialogue with “TERRITORIES Library group”, as will also be described in Section 2.3 of the report.

In order to gain a systematic overview of the potential processes in a given environment that may require attention from a modelling perspective, Features-Events and Processes (FEPs) can be employed. An example is provided in Section 3.1. Furthermore, a key component of determining the fitness-for-purposiveness of a given model requires the selection of a methodology for comparing correspondence between modelled and empirical datasets. This has required a comprehensive review of relevant literature from environmental sciences in general and deliberation as to the most appropriate method for radioecological type studies in particular. This subject is also covered in Section 3 with guidance provided on an approach for application in successive parts of this report.

The next stage of the approach involves the application of existing, widely applied models, the selection of which was informed by the experience of TERRITORIES modellers as noted above, to the various case study sites. Typically, an assessor/modeller would access previously developed approaches and tools when confronted by a given problem requiring the simulation of radionuclide behaviour and fate in the environment. Therefore, although the procedure presented was artificially contrived, it was intentionally devised to represent the reality of situations involving and challenges opposing the modelling process. A brief description of the initial models that were available to the TERRITORIES modelling groups are described in Section 4. Building on experience gained through the operation of existing, widely used models, it was possible to identify where further developmental work was necessary. This took the form, for example, of the application of bespoke models for specific sites that were under development by some of the TERRITORIES modelling groups independently from the analysis planned within this report or for specific cases where limitations in existing models were

² NORM = Naturally occurring radioactive material – primarily characterised by natural decay series (²²⁸U and ²³²Th decay chains) radionuclides.

evident. The focus of the second part of Section 4 has thus been placed on the description of more advanced (often, more process-based) models.

In Sections 5, the models ERICA TOOL / CROM, NORMALYSA and GRANIS have been applied at forest sites contaminated by NORM in Fen-Norway. At the Belgian NORM site, as considered in Section 6, the models: FORESTCROM, ECOFOR (ECOlogical model for FOrest Radioecology) and TALAL (Transport and Availability of elements and water between vegetation and soil) have been applied. For the Fukushima site (Section 7), as simple models we used the pre-existing IRSN's TREE4 model (included in SYMBIOSE) whilst CIEMAT and UT developed a model FORESTCROM based on IAEA SRS19 (IAEA, 2001) approach for forests. As advanced models to compare, an improved version of the TREE4 model (still not implemented in SYMBIOSE) was used. The performances of these models are evaluated with regards to field measurements of radiocesium (^{137}Cs) contamination in Japanese forests contaminated by the Fukushima-Daiichi NPP atmospheric fallouts. This dataset which is available in the TLD includes measurements of radiocesium activity concentrations, inventories and fluxes within the soil-tree system, over the period March 2011 – March 2017. Finally, the ARCTICMAR and AMIS models have been applied for the Sellafield site as described in Section 8. It was quickly observed that for this case, existing models could not be used to simulate the behaviour and fate of hot particles in the marine system. This instigated further model modification and development. In this part of the report (Sections 5 to 8), comparisons with empirical data from the TLD, drawing on the approaches presented in Section 3 where practicable, are made.

There is an attempt to comment, as far as practicable, with regards to how the development of new models has improved the capability of the simulation process. This has required a brief recap of the existing, widely applied models versus advanced models through a short discussion of results, limitations of approaches and quantification of performance as presented in Section 9. The plan was to consider good modelling practice with due attention to some of the typical methods employed such as implementation of tiered approach, in line with the original intention to structure such attempts 'in a graded way', and inclusion of quality assurance procedures. Stakeholder requirements were also determined to be of relevance and an attempt is made to link such perspectives to guidance on the theme in hand. For completeness, specific advice related to the consideration of uncertainty when implementing models is addressed in Appendix 12.2. This has been a subject of a CONCERT-TERRITORIES deliverable report D9.62 (Urso *et al.*, 2019). However, efforts are made to supplement this work with additional guidance within the context of defining a Fit-for-purpose level of complexity in model selection.

Finally, the lessons learnt, that have emerged from the process outlined above, along with more general considerations, arising from the experience of the report authors in relation to risk assessment, have allowed the provision of recommendations (Section 10) within the context of establishing a germane fit-for-purpose level of complexity for model development.

2.3 Expectation and purposes

In the TERRITORIES project, case studies have covered a set of radioactively contaminated sites or areas. This set, hereafter named the TERRITORIES Library, covers a large geographical scope (Europe and Japan), and a wide range of source terms (natural and artificial radionuclides), of ecosystems, of spatial extent and occupation features (from fenced sites to inhabited areas), of temporal scales of interest (with long-term series up to 3 decades after Chernobyl), and of remediation histories. Some of these case-studies are further developed in the present deliverable, cf. sections 5 to 8.

The present Section 2.3 is based on the wider set of case studies, including those developed for other TERRITORIES tasks, and presents the result of brainstorming with scientists involved in these case studies about the typical purposes and expectations regarding radiological characterization.

Received answers covered three types of operational expectations:

-for zoning, e.g.

- to identify zones that can be returned to normal use without any decontamination.
- to identify ‘hot-spots’ to be closed for public entrance. This specific purpose was for example mentioned in the context of the Semipalatinsk case study, developed in CONCERT-TERRITORIES deliverable report D9.65 (Guillevic *et al.*, 2018) to identify agricultural zones where to stop production, and to identify agricultural zones where to continue agriculture (or production), with or without counter-measures. During a stakeholder panel reported in CONCERT-TERRITORIES Deliverable report D9.66 (Maître *et al.*, 2019), wine producers confronted with an hypothetical post –accidental situation identified the need to have “information about the precise area affected by the radioactive deposits and its precise delimitation”, as it appeared to them “essential to avoid generalization and stigmatization” of a whole region

-for impact assessment, e.g.

- for prospective and retrospective assessments to human population. As an example, the specific question “Are doses to representative persons below relevant dose constraints and limits?” was identified for the case study about exposure on Cumbrian beaches, see Sellafield case study, section 8 in the present deliverable, and further development by CONCERT-TERRITORIES Deliverable report D9.63 (Jones *et al.*, 2019).
- for prospective and retrospective assessments to biota.

-for remediation, e.g

- to quantify waste that will, or could, be generated by a remediation option.
- to study the efficacy of a remediation option. As an example, the specific objective to investigate the sustainability and the efficiency of the radium trapping within NORM sediment of a freshwater reservoir was identified in the case study about the Rontok Lake in Poland, developed in CONCERT-TERRITORIES deliverable reports D9.59 (Smith *et al.*, 2019) and D9.60 (Mora *et al.*, 2019).

In addition to the three above-mentioned categories of operational purposes, some TERRITORIES case studies have also targeted more fundamental purposes, e.g. to gain mechanistic understanding of the processes involved in the transfer of a given element. In such a case, the radioecological model is intrinsically part of the research, and the criteria to evaluate if the model is “fit-for-purpose” are linked to the choice of processes to be characterized.

Regarding the operational purposes, the expectation was usually primarily expressed as a need for data, i.e. radioecological model outputs do not appear expected per se, but seem to be considered as an alternative to data, i.e. to cover gaps (e.g. if data are available for one compartment, but not for another one), or to predict the future. This perception of calculations as an alternative to data was for example explicit in CONCERT-TERRITORIES deliverable report D9.65 (Guillevic *et al.*, 2018), regarding

the progress of a post-accidental situation: “All these evaluations made during the emergency and transition phases are largely based on calculations, referenced data or sometimes assumptions (self-evacuation of population or not) and subjects to uncertainty in the prediction of the exposures of the population and the contamination of the environment. Complementary data based on monitoring (from the transition phase and later) are necessary to have an accurate knowledge of the situation (contamination of the environment, description of the landscape, population) and to support or adapt the first countermeasures decisions (zoning, restriction of consumption) already taken or initially planned during the previous phase.” For this reason, the ability of a model to reproduce data, when data are available for model validation, appears as a prerequisite for its acceptance. This explains why a key component for determining whether a given model is fit-for-purpose is the selection of a methodology for comparing correspondence between modelled and empirical datasets, as further developed in the present deliverable.

3 HOW TO IMPROVE A MODEL

As was mentioned in the previous section, the modelling of the behaviour of the different radionuclides in any contaminated territory can be improved by moving from empirical models to process-based models. The process-based models consist of objects or entities and processes that characterized internal structure of these objects and their interaction. So, they should ground on a good theoretical understanding of an ecological system which will modelled and accounting of all relevant processes related with this system.

One of the ways to take into account all the relevant processes is the method of IM (interaction matrices) and FEP (Feature, Events, and Processes) analysis.

3.1 Interaction matrices – FEPs analysis

3.1.1 Principles

The natural ecosystems (e.g. forest ecosystem) involve multiple components and processes during the transfer of radionuclides through these components. To take into account and congruently organize all relevant components and processes the interaction matrices can be used.

Interaction matrices, an expert semi quantitative method to identify multiple interactions between biotic and abiotic components of the ecosystem can be considered as a useful tool to develop conceptual models of the behaviour of radionuclides in the environment. This systematic approach facilitates the comprehensible identification of the pathways of the main radionuclides and permits classification of the role of different ecosystem components in terms of cause-effect relationships.

The interaction matrices methodology is based on the description of the components (compartments) and their interactions using a $n \times n$ matrix, X_{ij} , where i – is the row number and j – is the column number. In these matrices the diagonal elements ($i = j$) are compartments and non-diagonal elements ($i \neq j$) are interactions between them (Avila *et al.*, 1999; Velasco *et al.*, 2006). Elements of the matrix with $i < j$ show how compartment X_{ii} influences the compartment X_{jj} (blue arrows in Figure 3.1-1) and elements with $i > j$ show how the compartment X_{ii} is influenced by X_{jj} (green arrows in the figure). These influences are named as features, events and processes (FEPs). This includes physical features, events and processes that could directly or indirectly influence the release and transport of radioactive elements in the ecosystem, plus other factors like regulatory requirements, that constrain or focus the analysis.

Features, Events, and Processes analysis (FEPs) was developed in the nuclear waste area (Chapman *et al.*, 1995), and is also proposed for other problems like storage of CO₂ (Savage *et al.*, 2004). In the FEPs analysis 'Features' mean physical, chemical and other characteristics of the systems, 'Events' mean the phenomena, which can occurred in the systems with some probability and 'Processes' mean the inherent processes occurred in the systems.

FEPs analysis can also be applied for study the evolution of contaminated ecosystems e.g. forest ecosystems and fate of radioactive contaminants in them. In this context 'Features' mean physical, chemical and other characteristics of the forest ecosystem (e.g. type of trees, soil etc.). 'Events' mean the occurrences that have a specific starting time and, usually, duration shorter than the time being simulated in a model, e.g. fire in contaminated forest, floods etc. 'Processes' are the physical, chemical and other processes and interactions in the different components of the forest ecosystem which were described more detailed in the previous section.

The interaction matrix approach is a systematic approach that shows how FEPs are incorporated in the conceptual modelling process.

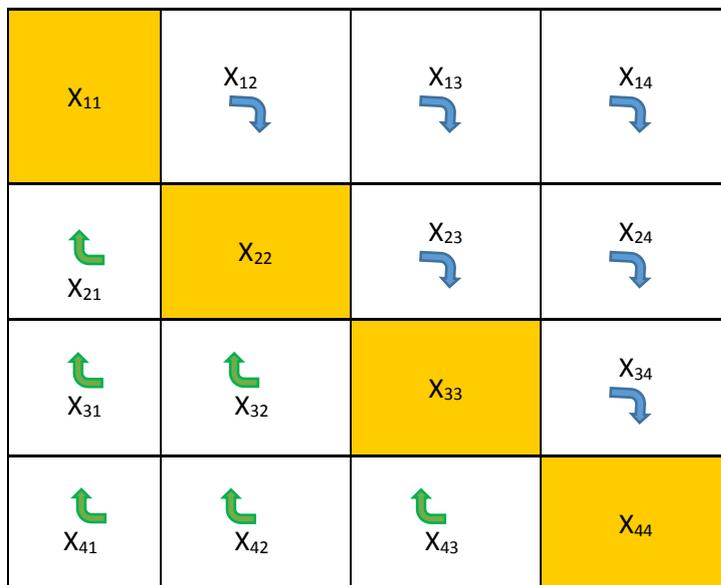


Figure 3.1-1: Schematic example of a 4x4 interaction matrix.

The degree of interaction between components can be scored, e.g. from 0 to 1, where 0 means that interaction doesn't exist and 1 means that strong interaction exists. The assignment of such scores is subjective and depends on a particular expert opinion. The involvement of several experts in the process of scoring can increase the objectivity in this process. Such work was done within the STAR project (Sweeck *et al.*, 2014) as an example for prioritizing processes in interaction matrices using a forest scenario from Chernobyl exclusion zone (see next section).

Non-diagonal components of such matrices are elements of mathematical models that characterized the behaviour and fate of radionuclides within a given ecosystem.

These matrices can help to build the conceptual model and then the math model of systems of differential equations, which describe the processes reflected as interactions between the compartments:

$$\frac{dX_{ii}}{dt} = \sum k_{ji} X_{jj} - \sum k_{ij} X_{ii} - \lambda_r X_{ii} \quad (3.1)$$

Being

X_{ii}, X_{jj} - Concentration of the activity of the radionuclide in the compartment ii, jj (soil, tree, litter etc), Bq/m²; Bq/kg (dependent on the compartment);

k_{ij} - is the transfer coefficient of the radionuclide from compartment ii to compartment jj ; d⁻¹

λ_r - is the decay rate of radionuclide r , d⁻¹.

In the TERRITORIES project the evolution of forest ecosystems as Fukushima, Norwegian fen forest site or Belgian NORM site, are studied based in the natural evolution of already existing radioactive contamination without taking into account possible external phenomena as fires, floods etc.

In the following sections of this report the application of process-based models for selected sites are demonstrated taking into account all relevant processes for these sites.

An example of an interaction matrix developed for a forest in the STAR project is shown below.

3.1.2 Example: Interaction matrix for a forest ecosystem

A prioritization exercise carried out in the project STAR (Sweeck *et al.*, 2014) by several scientists from 7 European institutions (BFS, CEH, CIEMAT, IRSN, STUK, SU, and NRPA). To analyse interactions in the forest ecosystem the European Observatory site for Radioecology of the Chernobyl exclusion zone -, a zone with a significant contamination of artificial radionuclides (mainly ¹³⁷Cs and ⁹⁰Sr) was used.

Such compartments of the diagonal matrix of the forest system have been considered:

1. Atmosphere
2. Groundwater
3. Deep soil
4. Root zone/active layer
5. Litter
6. Decomposers
7. PP natural understory
8. PP natural canopy
9. PP agricultural
10. PC natural
11. PC domestic
12. SC natural
13. Humans

Where PP means 'Primary producers', PC – 'Primary consumers' and SC – 'Secondary consumers'

The main objective of the exercise was to prioritize the FEPs for the development and simplification of a conceptual model for a given forest. As a secondary objective, the possibility of using the results of the prioritization as a tool for prioritizing the radioecological research on the Observatory sites for Radioecology was raised.

Following rules were considered to develop the interaction matrix:

- A process described by all the partners as important was considered ‘Main Common Process’ (see list below)
- A process where all partners except 1 had described a process as important and the single partner considered it as neutral, was considered a ‘Secondary Process’ (see list below)
- Cells where all partners categorised processes as “neutral” or “non-important” were not considered in the interaction matrix.

The next list of main common processes for the forest was defined in Table 3.1-1 and interaction matrix of forest ecosystem is presented in Figure 3.1-2.

Table 3.1-1: Main processes in forest ecosystem selected in STAR project.

Processes	Compartment which influences	Compartment are influenced by
Ingestion, external exposure, inhalation	root zone	decomposers
root uptake, respiration, foliar transfer, external exposure, photosynthesis	root zone	PP natural understory, PP natural canopy, PP agricultural
ingestion, inhalation, external exposure, adsorption	root zone	PC natural, PC domestic, SC natural and Humans
decomposition	litter	root zone
ingestion, external exposure, inhalation	litter	decomposers
root uptake, respiration, foliar transfer, external exposure, photosynthesis	litter	PP natural understory, PP agricultural
ingestion, inhalation, external exposure, adsorption	litter	PC natural, SC natural, PC domestic, SC domestic
inhalation, external exposure,	litter	Humans
Ingestion	Decomposers (fungi)	Humans
ingestion, external exposure	PP natural understory	PC natural, PC domestic, Humans
ingestion	PC natural	SC natural, Humans
ingestion	PC domestic	Humans

Root zone/ active layer		Ingestion; ext. exposure; inhalation	root uptake; respiration; foliar transfer; ext.exposure; photosynthesis	root uptake; respiration; foliar transfer; ext. exposure; photosynthesis		Ingestion; inhalation; ext. exposure; adsorption		Ingestion; inhalation; ext. exposure; adsorption	Ingestion; inhalation; ext. exposure; adsorption
decomposition	Litter	Ingestion, external exposure; inhalation	root uptake; respiration; foliar transfer; ext. exposure; photosynthesis			Ingestion; inhalation; external exposure; adsorption		Ingestion; inhalation; ext. exposure; adsorption	Inhalation; ext. exposure; adsorption
		Decomposers	symbiotic uptake	symbiotic uptake		ingestion	ingestion		ingestion
			PP natural understory			Ingestion; ext. exposure			Ingestion; ext.exposure
				PP natural canopy					
					PP agricultural		ingestion ext.exposure		Ingestion; ext.exposure
						PC natural		ingestion	ingestion
							PC domestic		
								SC natural	
									Humans

Figure 3.1-2: Interaction Matrix of forest ecosystem selected in STAR project

3.2 Need for a methodology to quantify model improvement in radioecology

The quantification of model improvement is useful to assess whether an ‘advanced’ model compared to a ‘simple’ one provides a model output with reduced overall uncertainty. Quantifying model improvement means quantifying differences in model performance and accuracy of model output. This is achieved by using experimental data for comparison with model output and/or by making use of qualitative criteria, especially when data are unavailable or when forecast predictions are needed.

Improvements in models are undertaken for several reasons:

First, whenever an increased understanding of any physical, chemical or biological process is available, it may be worth to include it (see in Figure 3.2-1 the processes involved in the transfer of radionuclides through the different ecosystem compartments in a forest). In this case, it is of interest to check whether this translates into better model performance, i.e. the difference between model output and experimental data is reduced.

In radioecology, for example, significant effort is made to improve models that describe radionuclide uptake by plants. Often the so-called soil-plant concentration ratios (CR) are used, which are location-specific and can vary for a given radionuclide over orders of magnitude. CRs relate the radionuclide concentration in plant (Bq kg^{-1} dry weight or fresh weight) to that of the soil (Bq kg^{-1} dry weight) in the form:

$$\text{CR} = \frac{\text{Concentration in plant (Bq kg}^{-1}\text{dw or fw)}}{\text{Concentration in soil (Bq kg}^{-1}\text{dw)}} \quad (3.2)$$

CRs are an example of a ‘simple’ model in which all processes are considered in an aggregated way. In contrast, a model that provides as output the radionuclide concentration in a specific plant component by considering discretized soil profiles (as shown in Figure 3.2-1) is an example of an ‘advanced’ model.

Second, models may need to be adapted whenever new or additional requirements have to be considered. Adaptation to new requirements often implies that the overall uncertainty of the model is quantifiable and that envisaged improvements reduce it (Urso *et al.*, 2019). As an example, the German General Administrative Regulation for Practices is under review in light of the new requirements in the EURATOM basic safety standards (Council Directive 2013/59/Euratom) (Official Journal of the European Union, 2014): according to these new requirements, the dose calculated for members of the public needs to be as realistic as reasonably possible. This implies that the overall uncertainty of the assessment model used has to be as low as reasonably possible. In addition, the model must be suitable for demonstrating compliance with dose limits. In order to ensure that the dose limits are not exceeded, the model has to include an adequate degree of conservatism that reflects the overall uncertainty of the model. This is another argument for assessment models with a small uncertainty budget.

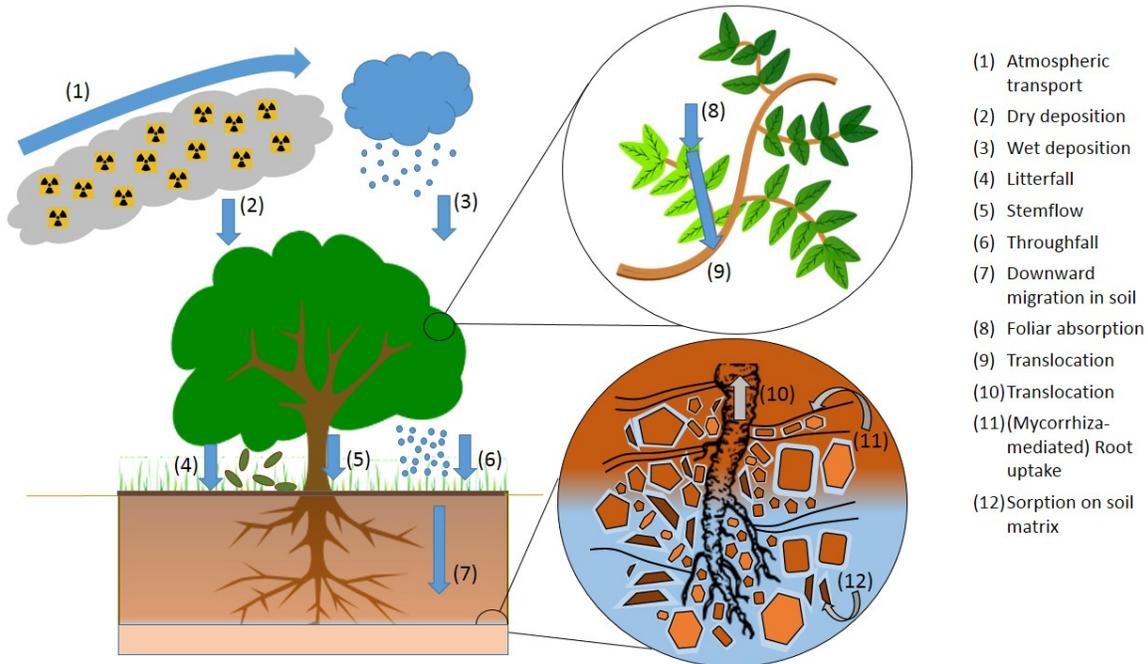


Figure 3.2-1: Visualisation of the processes involved in the transfer of radionuclides through the different ecosystem compartments in a forest.

The development of a general methodology to quantify the improvement of radioecological models is meant to support the reader through the critical analysis of approaches to assess model performance, thus assisting in choosing among different tools and metrics available for comparing model outputs. The aim of this work is to provide the reader with a simple methodology to compare ‘simple’ and ‘advanced’ models in radioecology by making use of available tools from other disciplines and by applying them to radioecological models to test their performance. In this report, it is assumed that the models considered have been calibrated and validated with an adequate data set, before assessing model performance.

3.3 Model performance

3.3.1 What does model performance mean?

In radioecology, two types of models can be distinguished: research models and assessment models.

Research models rely on a detailed understanding of the processes and mechanisms responsible for the transfer of radionuclides in the environment. They can have a complex mathematical structure and their focus is on the reproduction of measured data in order to confirm proper understanding of the processes involved. These models can have a large number of parameters and need to be calibrated and validated with site-specific data. However, they are the basis for deriving simplified, scientifically based assessment models.

Assessment models focus on the prediction and forecast of activity levels (AL) and ambient dose rates, e.g. for regulatory purposes, for forecasting the future evolution of contamination levels, for obtaining information on contamination patterns. They tend to have simpler mathematical structures than research models.

Environmental processes and aggregated model parameters are often characterised by large spatial and temporal variability. In fact, concentrations in any environmental medium can vary over orders of magnitude at a specific contamination site. On the other hand, experimental data are often scarce or even lacking, because either concentration levels are below the detection limit, or the number of samples that could reasonably be measured is too small.

In radioecological assessment models, the ultimate quantity (endpoint) is the dose to humans and non-human biota. In this case, the dose calculated from the model output can be compared with the dose calculated from measured data. As already mentioned, assessment models for regulatory purposes often require that the quantification of the dose must be conservative, i.e. the calculated dose must not underestimate the real one.

Model performance in general is a measure of how fit a model is for its specific purpose. In other words, the definition of a 'good' model depends on the application and need; it is virtually impossible to find a model that is universally 'good' for all applications and needs. In fact, a model has a set of properties that make it suitable (or not) to investigate specific aspects of an assessment task. As an example, if one requires a quick estimate of activity levels in the environment after an accidental release of radionuclides, a complex and data-hungry model would not be adequate for a first, quick assessment, as it may take a longer time to provide results. However, if an assessment is needed to decide the type and extent of remediation activities in a contaminated area, it may be worth to use models that can reproduce spatial and temporal heterogeneity in detail, in order to minimise the economic effort of such a remediation activity.

A model for assessing the order of magnitude of contamination levels to be expected in biomass is the so-called aggregated transfer factor T_{agg} . Such a model is fit for purpose whenever activity levels have to be estimated very quickly, e.g. after an accidental release of radionuclides. The T_{agg} values are calculated as follows:

$$T_{agg} = \frac{\text{Activity concentration in biomass (Bq kg}^{-1} \text{ dw or fw)}}{\text{Inventory in soil (Bq m}^{-2}\text{)}} \quad (3.3)$$

They refer to the total deposition on the ground. For long-term predictions of activity levels, however, more sophisticated approaches are necessary. As an example, the model developed by Rühm *et al.* (1998) for the time-dependent contamination levels of fungal fruit bodies is considered. In their approach, the concentration ratio for a specific fungal species (which is constant in time) refers to the radiocaesium concentration in the horizons of forest soil that the mycelium exploits. When coupled with a model for the migration of radiocaesium in soil, such an approach describes the measured contamination levels of the mushrooms satisfyingly. For the depth localisation of the mycelium in the forest soil, the isotopic ratio of ^{137}Cs and ^{134}Cs was used.

3.3.2 How to evaluate model performance

Model performance can be assessed qualitatively and/or quantitatively. A qualitative assessment is always possible and it is the only possibility if experimental data are not available or the forecast performance of a model has to be assessed. In this case, a set of qualitative indicators is the basis to judge if the model is adequate with respect to its purpose.

Quantitative evaluation of model performance is based on numerical measures (metrics) that can be calculated if experimental data are available. These metrics are described in Section 3.4. It is important to stress that different approaches and numerous metrics are available. Each of which exhibits pros and cons and can highlight only some specific characteristics of the model. The various available metrics need to be judged in terms of:

1. Ease of use and straight-forward applicability;
2. Acceptance and use within the broader scientific community, i.e. recommendations to use them, availability of examples in published literature, potential for benchmarking with other published results;
3. Identified strengths (and recognised weaknesses) in model evaluation.

3.4 Overview of numerical tools for evaluating model performance in environmental sciences

In the work of Bennett *et al.* (2013), numerous metrics are presented and categorised based on the requirements a model has to fulfil to be considered adequate for a specific purpose. The categorisation used in Bennet *et al.* (2013) is the following:

1. Direct value comparison
2. Coupling real and modelled values
3. Preserving the data pattern
4. Indirect metrics based on parameter values
5. Data transformation methods

Direct value comparison (category 1) is the simplest approach that can be used for quantification of model improvement. The quantified information focuses on whether the model output lies within a certain range of the available experimental domain. This approach is useful as a starting point to understand how model output and available data are linked. In radioecology, measurements often are characterised by high variability and heterogeneity and summary measures over a spatially and/or temporally broad range of values do not provide detailed information. An example of such heterogeneous data could be radionuclide concentrations in soil or in biota.

Coupling real and modelled values (category 2) is a more sophisticated approach, which provides quantitative assessments based on the point-to-point relation between model output and experimental data.

Preserving the data pattern (category 3) deals with the performance analysis of dynamic models. If data are available in form of time series, performance assessment focuses on the ability of the model to describe the experimentally observed changes over time.

Indirect metrics based on parameter values (category 4) allow to determine the relative ranking between models, but cannot be used to evaluate model performance in absolute terms, i.e. they are less relevant to assess how well the model approximates the data. The Akaike Information Criterion (AIC) (Burnham and Anderson, 2002), the Bayesian Information Criteria (BIC) (Schwarz Gideon, 1978) or the Deviance Information Criterion (DIC) (Spiegelhalter *et al.*, 2002) are examples of this approach. Advantage of the indirect metrics is that they provide a dimensionless measure of performance and include in the metric

calculation also the number of model parameters. However, they require the definition of probability distribution functions for the priors and the likelihood function. Therefore, results will be strongly affected by the choice of the probability distribution functions. In addition, indirect metrics tend to be computationally demanding.

Data transformation methods (category 5) are a technique that can be used to highlight aspects of the behaviour of a model that are not clear in the original time or space domain. An example is the use of Fourier transforms to isolate seasonality effects from yearly or daily effects.

3.4.1 Direct value comparison

The aim of direct value comparison is to test whether the model output y shows similar characteristics as a whole set of comparison data \hat{y} . An example is the standard summary statistics of both y and \hat{y} by means of calculating mean value, median, range or variance.

3.4.2 Coupling real and modelled values

These methods consider the pairs of values y_i and \hat{y}_i for the same point in time and in space i . The most commonly used methods are:

- Concurrent comparison analysis
 - Scatter plots
 - Linear regression analysis, where the slope and intercept are usually examined under the assumption that measured and simulated values are linearly related. It implies that all of the uncertainty variance is contained in simulated values and that measured data are free of uncertainty (Willmott, 1981).
 - In reality, measured data are rarely, if ever, free of uncertainty. Harmel *et al.* (2006) showed that substantial uncertainty in reported water quality data can result when individual uncertainties from all procedural data collection categories are considered. Therefore, care needs to be taken when using regression statistics for model evaluation.
 - Contingency table (Ghelli and Ebert, 2008), which is based on reporting model behaviours in important cases, for example, when passing a specified threshold. In the table, the number of occurrences in which real data and model output are above the threshold are reported.
- Residual methods like residual plots, quantile-quantile (QQ) plots (Wilk and Gnanadesikan, 1968), mean squared error (MSE) (Chai and Draxler, 2014), residual autocorrelation analysis and residual methods with data transformations.

The residual plot has the residual values on the vertical axis; the horizontal axis displays the independent variable. Ideally, residual values should be equally and randomly spaced around the horizontal axis. If this is not the case, the data set is not a good candidate for a linear regression and potential reasons must be investigated (e.g. missing variables, missing higher-order terms that explain a non-linear pattern).

The QQ plot tests whether the distribution of residuals approximates normality. Its statistical significance can be assessed with the Kolmogorov-Smirnov test (Chakravarti *et al.*, 1967).

Common residual methods are:

- Mean absolute error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.4)$$

- Mean squared error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.5)$$

- Root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.6)$$

where y_i are the modelled data, \hat{y}_i are the observed data and n is the number of data points available. Values of zero for RMSE, MAE and MSE indicate a perfect match. Singh *et al.* (2004) state that RMSE and MAE values less than half of the standard deviation of the observed data may be considered low and that either metric is appropriate for model evaluation.

The metric bias is the mean of the residuals, namely:

$$BIAS = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (3.7)$$

It provides information on whether the model tends to underestimate or overestimate the observed data, with an ideal value of zero.

The air pollution transport modelling community (Chang and Hanna, 2005) suggests the geometric mean bias:

$$MG = \exp\left(\frac{1}{n} \sum_{i=1}^n \ln \hat{y}_i - \frac{1}{n} \sum_{i=1}^n \ln y_i\right) \quad (3.8)$$

where y_i are the modelled data, \hat{y}_i are the observed data and n is the number of data points.

For deeper insight, it is recommended to split the MG (eq. 3.8) into underestimating and overestimating (or false negative and false positive) parts, denoted as MG_{FN} and MG_{FP} , respectively:

$$MG_{FN} = \exp\left[\frac{1}{2n} \cdot \sum_{i=1}^n [|\ln \hat{y}_i - \ln y_i| + (\ln \hat{y}_i - \ln y_i)]\right] \quad (3.9)$$

$$MG_{FP} = \exp\left[\frac{1}{2n} \cdot \sum_{i=1}^n [|\ln \hat{y}_i - \ln y_i| + (\ln y_i - \ln \hat{y}_i)]\right] \quad (3.10)$$

Thus

$$MG = \frac{MG_{FN}}{MG_{FP}} \quad (3.11)$$

The MG takes into account that model predictions and observed data may differ by a factor of ten or even more in radioecology. Linear measures give much higher weight to the overestimation than to the underestimation of data by the same factor. The logarithmic scale balances the weights of overestimation and underestimation.

3.4.3 Dealing with time series

If data in form of time series are available, it may be of relevance to test the ability of the model to preserve the temporal or spatial pattern of data. For this purpose, cross-correlation techniques and correlation coefficients (e.g. in hydrology the Nash-Sutcliffe efficiency coefficient NSE or the coefficient of determination R^2) can be applied. The primary goal of this type of performance testing is to understand how data points and their uncertainties relate to each other.

Given two functions f and g , f^* is the complex conjugate of f and τ is the displacement in time (also called lag), the general expression of the cross-correlation function is:

$$(f^*g)[\tau] = \int_{-\infty}^{\infty} f^*(t)g(t + \tau)dt \quad (3.12)$$

The cross-correlation function can be calculated between observed and modelled values. In this case, the cross-correlation function (for real quantities) can be expressed as:

$$CCF = \sum_{t=1}^{\infty} y_t(y_{t+\tau} - \hat{y}_{t+\tau}) \quad (3.13)$$

where τ is the lag considered, y_t is the model output, \hat{y}_t is the data set and $(y_{t+\tau} - \hat{y}_{t+\tau})$ is the residual between model output and data, given that data are in form of a time series $(\hat{y}_{t1}, \hat{y}_{t2}, \hat{y}_{t3}, \dots)$.

Cross-correlation can be calculated between input data u_t and residuals as well, namely by using the function:

$$CCF = \sum_{t=1}^{\infty} u_t(y_{t+\tau} - \hat{y}_{t+\tau}) \quad (3.14)$$

where u_t is the input quantity of the model (e.g. rainfall in mm/day) and $(y_{t+\tau} - \hat{y}_{t+\tau})$ is the residual between model output and experimental value. In this case, significant correlation between input data and residuals may indicate unmodelled behaviour.

The so-called Pearson product moment correlation coefficient (PPMCC) (Ross, 2017) calculates the correlation between two series of sampled data:

$$PPMC = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (3.15)$$

where \bar{y} is the mean value of the modelled data and $\bar{\hat{y}}$ is the mean value of the observed data. PPMC ranges from -1 to 1 and is a measure of the degree of linear relationship between observed and simulated data. If $PPMC = 0$, no linear relationship exists. If $PPMC = 1$ or -1 , a perfect positive or negative linear relationship exists.

The coefficient of determination (R^2) (Boddy and Smith, 2009) is a squared version of the PPMC and it varies between 0 and 1 . Similarly to the PPMC, R^2 describes the proportion of the variance in observed

data explained by the model. R^2 ranges from 0 to 1, with higher values indicating less error variance, and typically values greater than 0.5 are considered acceptable (Santhi *et al.*, 2001; Van Liew *et al.*, 2003).

The Pearson product moment correlation coefficient (PPMC) and the coefficient of determination (R^2) describe the degree of collinearity between simulated and observed data. Although PPMC and R^2 have been widely used for model evaluation, these statistics are over-sensitive to high extreme values (outliers) and insensitive to additive and proportional differences between model predictions and observed data (Legates and McCabe, 1999).

Moriasi *et al.* (2007) provide a categorisation based on tools applied in watershed model analysis. Moriasi *et al.* (2007) categorise the different metrics into three groups:

1. Standard regression statistics that considers the strength of linear relationship and the coefficient of determination R^2 .
2. Dimensionless metric that allows for straightforward comparison among models, e.g. Nash-Sutcliffe efficiency coefficient or bias expressed as a percentage.
3. Error index metric that deals with deviations in the units of the data of interest, e.g. root mean squared error (RMSE).

Moriasi *et al.* (2007) highlight the importance of relying on the experimental data available if a quantitative assessment of model performance is the aim.

Analysis of model outputs can benefit from updates via data assimilation and Bayesian statistics if new data are available (Reich and Cotter, 2015). It is worth to mention that in radioecology analysis of model performance with Bayesian statistics is used for example in Sy *et al.* (2015) by considering the posterior predictive loss criterion (PPLC) from Gelfand and Ghosh (1998) as a suitable metric. The PPLC method is a quantitative criterion that measures the gap between the observations and the corresponding predictions obtained. In Sy *et al.* (2015) it has been combined with the Bayesian statistics. However, this is an advanced method and will not be considered further in this report.

3.5 Suggested metrics for radioecology

For testing the performance of radioecological models, we will focus on the RMSE (eq. 3.6). Its use is motivated by its close analogy to the reduced chi-square (χ_{red}^2) and the underlying theory, i.e. the maximum likelihood estimation method (Barford, 1967).

The RMSE is the most popular evaluation metric used in regression problems. Its use is recommended in literature (see Bennet *et al.* 2013; Moriasi *et al.* 2007). It is based on the assumption that errors are unbiased and follow a normal distribution.

The key points to consider for RMSE are the following:

1. The 'squared' nature of this metric prevents the cancellation of positive and negative error values and thus delivers more robust results. In other words, this metric aptly displays the plausible magnitude of error term.
2. RMSE has the disadvantage to be highly affected by outlier values. Hence, one has to ensure that outliers have been removed from the data set prior to using this metric.

3. As compared to mean absolute error (MAE), RMSE gives higher weights to large errors.

In addition to the RMSE metric, the bias metric, the cross-correlation functions and the Pearson product moment correlation coefficient could also be considered. The bias metric (eq. 3.7) can be used to check to what extent the model output underestimates or overestimates the data. The cross-correlation functions (eqs. 3.13 and 3.14) and/or the Pearson product moment correlation coefficient (eq. 3.15) can be used to check whether the model output can properly preserve the time pattern of data in case time series data are considered.

The metrics suggested above are ‘standard’ metrics, partly already used in radioecology (e.g. RMSE), which fulfil the judgement criteria 1 to 3 in Section 3.2. For these metrics, enough information and experience summarised in published material is available. Nevertheless, if the reader is interested in more advanced methods, he/she can refer to Bennett *et al.* (2013) for examples (e.g. indirect metrics and data transformation methods) and further reading.

In radioecology, a modified version of RMSE is most suitable, taking the logarithm of model results and observed data. Details are given in Section 3.6.

3.6 Availability of experimental data

In order to quantify model improvement, the performance of each model considered has to be tested. This can be quantitatively achieved by comparing model output with experimental data. In this way, model performance is also a measure of model accuracy. Experimental data need to be available and adequate to perform the analysis: uncertainties have to be known and available data have to be representative of the location considered.

In general, data representativeness refers to a phenomenon where the collected data (a sample) accurately reflect the question under study, e.g. a specific environmental process. Ramsey and Hewitt (2005) note that assessment of data representativeness is only possible after clearly stipulating the targeted population and the purpose of the assessment. Having a large sample does not imply representativeness; rather the manner in which the sample was collected plays an important role in ensuring representativeness. If the selection of the sample is biased towards elements that either have the desired characteristic of interest or have similar characteristics, then even a generously large sample will not deliver representative data. Process characteristics estimated from such data will be biased towards the preferred elements.

The endpoints of radioecological models, especially assessment models, may differ from the type of available data. In this case, two approaches should be considered:

1. Use only the data sets where model output and measured data match.
2. Consider ‘interim results’ of the model that match with the measured data available (e.g. activity levels in environmental media).

However, care should be undertaken with option 2, since time scales and spatial scales may differ between interim results and model output.

3.6.1 Methods for data splitting

The data available have to be used not only to test the model performance but initially also to calibrate and validate the model. This implies that proper methods need to be considered to split the data for calibration, validation and performance assessment.

Several well-established methods can be used to split data, namely:

- Cross-validation methods
 - Hold out method (Kohavi, 1995), in which the data are split into two groups, one for development and one for testing. The sizes of both groups and the way of group splitting affect the apparent performance of the model and the accuracy of the testing.
 - K-fold partitioning (Kohavi, 1995), in which data are split into k sets, one set is used for training, the remaining k-1 sets for testing. The method is then repeated k times and the results are averaged.
 - Leave-one-out cross validation (LOOCV) (Kohavi, 1995; Hastie *et al.*, 2009), in which n-1 data points are used for model development and only one point is used for validation. This is repeated for all data points, each one in turn being left out.
- Bootstrapping (Efron and Tibishirani, 1993), in which the set of data is randomly re-sampled with replacement of the original measurements. This is repeated multiple times to estimate the error distribution of the model. This method usually requires a large number of data points in order to obtain independently distributed residuals.

It is not possible to provide a rule of thumb to determine the minimum number of data points necessary for applying such data splitting methods. In literature, such methods are used for data sets that include hundreds of data points (e.g. Gonze and Sy, 2016).

3.6.2 Model performance and model uncertainties

When data are available to quantitatively assess model performance, the difference between model output and measured value (the residual) is assumed to be the overall (total) uncertainty of the model. Ideally, the metric for estimating model performance should take into account the uncertainties of measured data, conceptual model uncertainty, parameter uncertainty and scenario uncertainty (Urso *et al.*, 2019).

When evaluating different models using a single metric, differences in the error distributions become more important. The MAE (eq. 3.4), for example, is suitable to describe uniformly distributed errors.

In the RMSE metric, it is implicitly assumed that errors are normally distributed. It is worth to consider the analogy with the reduced chi-square (χ_{red}^2):

$$\chi_{\text{red}}^2 = \frac{1}{n-1} \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{\sigma_i^2} \quad (3.16)$$

where n is the number of data considered, y_i is the modelled value at location or time i, \hat{y}_i is the respective observed value and σ_i^2 is the variance that includes the overall uncertainty of the model output and the

uncertainty of the data. Ideally, χ_{red}^2 should be close to 1 ($\chi_{\text{red}}^2 \approx 1$). $\chi_{\text{red}}^2 \gg 1$ indicates that the model output poorly agrees with the observed data with respect to the variances σ_i^2 , i.e. the variances σ_i^2 are too small.

In addition, one should consider that model results and observed data could match just by chance (random agreement). Random agreement has been quantified in the past in terms of the so-called kappa statistics (Kuhnert *et al.*, 2006). Newer methods to cope with this are summarised in Bennet *et al.* (2013).

3.6.3 Proposed methodology in radioecology

In this work, a methodology for radioecological models is proposed based on the combination of both a list of qualitative criteria for qualitative assessment and a modified form of the RMSE, namely its logarithmic version, for quantitative assessment.

Combining several qualitative and quantitative criteria aims at gathering a complete overview of what is feasible considering the purpose of the model, the typical characteristics of a radioecological model, and the amount and type of data available.

3.6.4 Description of suggested qualitative criteria

The suggested qualitative criteria should take into account that a comparison of modelled and measured data provides no information on the quality of forecasts, only on the reproduction of measured data. The following qualitative criteria can be used:

1. If dominant processes and mechanisms that determine the fate of radionuclides in the environment are modelled in a robust way, then the quality of model forecasts could be considered as acceptable, even if data to validate the model are unavailable. Robust modelling requires, for example, a reasonably small number of time-independent model parameters.
2. When a model is successfully applied to a data set from the past, and data for the present situation are not available, it could also be considered as acceptable. An example of this is the information obtained from the fall-out of nuclear tests on the surface in the 1950s and 1960s to track the migration of radionuclides in soil soon after the Chernobyl accident.
3. If measured data are not available for the radionuclide but are available for a stable isotope of the same chemical element and/or a chemical analogue, models that have been used for these analogues could be accepted for the radionuclide if the physical decay is taken into account. There should be reasonable confidence that the processes and mechanisms, e.g. the availability for plant uptake, are the same for the radionuclide and the stable isotope.
4. If measured data are unavailable at present or from the past, a model that accurately accounts for the dominant processes in detail is to be judged better than a simple one.

An example for neglecting an important mechanism is the use of the transfer factor model in order to quantify the natural ^{40}K level in humans. In contrast to the predictions of this model, the level of ^{40}K in humans does not depend linearly on the ^{40}K intake. Since potassium is homeostatically regulated in humans and natural potassium occurs with a fixed ratio of ^{40}K to total potassium, the ^{40}K level in humans is fairly constant.

The specific activity model for ^{14}C (Wirth, 1982) is an example that simple models may be fit for purpose. The model states that environmental processes affect all carbon isotopes in the same way, apart from some minor discrimination effects. Hence, the activity ratio $^{14}\text{CO}_2$ to $^{12}\text{CO}_2$, in ambient air is, to a good approximation, the same as the isotopic ratio $^{14}\text{C}/^{12}\text{C}$ in plants, animals and humans. Without modelling the dominant processes in detail, the specific activity model for ^{14}C proved to be very successful under equilibrium conditions. The concept is fundamental for carbon dating.

3.6.5 Description of suggested numerical measure

When dealing with radioecological models, a specific feature that has to be accounted for is the potentially huge variability of measurable endpoints, i.e. contamination levels of environmental media and biota as well as ambient dose rates.

In case data are very sparse and variations of orders of magnitude are involved, the logarithm of the data and the model output may be used in the RMSE metric, which increases the weighting towards small values. This metric is the so-called root mean squared logarithmic error (RMSLE):

$$\text{RMSLE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log y_i - \log \hat{y}_i)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(y_i / \hat{y}_i))^2} \quad (3.17)$$

Whereas the RMSE (eq. 3.6) quantifies the absolute difference between model output and measured data, the RMSLE (eq. 3.17) quantifies the ratio of model output and measured data, thus balancing the weights of small and high values. In addition, the RMSLE is dimensionless. In this way, one has the advantage of not having to change units whenever two models with outputs in different units are compared (for example if one model delivers an output in Curie and the other model delivers an output in Becquerel).

In radioecology, there are already examples of the use of the RMSLE, see e.g. Sy *et al.* (2015); Gonze and Sy (2016).

3.6.6 Performance table for combining qualitative and quantitative indicators

Ultimately, it is convenient to combine both qualitative and quantitative indicators in a performance table (Table 3.6-1), in order to provide a simple scheme for model performance assessment.

Table 3.6-1: Performance table for quantifying performance of a radioecological model

	Properties	Limitations
Qualitative performance assessment	Complements quantitative indicators; Provides indication about acceptability of a model in case data are not available to calculate any quantitative indicator.	Strongly relies on expert judgement.
Number of parameters involved and their time-independence	Provides a hint concerning over-fitting and indicates if parameters are robust for forecasts.	Depends on the type of model (assessment model, research model).

	<u>Properties</u>	<u>Limitations</u>
Number of processes considered separately	Provides a hint concerning the level of detail of a model and the understanding of the processes involved.	Depends on the type of model (assessment model, research model).
Application of the model to past set of data	Demonstrates the reliability of a model.	Cannot guarantee that the model works well for present data and for forecasts.
Application of the model to stable isotopes and chemical analogues	Possible approach when data for the radioactive isotope are not available.	Need to include physical decay; Requires reasonable confidence that the processes and mechanisms are the same for the radionuclide and the stable isotope.

	<u>Properties</u>	<u>Limitations</u>
<u>Quantitative performance assessment</u>	Makes use of available experimental data.	Provides a measure of objectivity; Need to check that data are representative for the given assessment purpose.
Root mean squared logarithmic error (RMSLE)	Takes care of the large variability of radioecological data; Dimensionless; Related to maximum likelihood method.	Need to verify whether there are outliers in the set of data, which should be excluded.
BIAS	Provides information on underestimation and overestimation of the data.	Result of zero does not indicate low error due to potential cancellation effects.
Geometric mean bias (MG)	Provides balanced information on underestimation and overestimation of data.	Interpretation not straightforward as for linear bias.
Cross-correlation function	Provides information on whether the model can preserve data pattern of time series.	Need to analyse not only cross-correlation between observed and calculated values but also cross-correlation between data and residuals.
Advanced tools: Indirect metrics	Weight the number of parameters; Allow for direct inter-comparison among models; Dimensionless.	Most of them require Bayesian/probabilistic techniques, i.e. knowledge of probability distribution functions for priors and likelihood is necessary; Large computational effort.

The methodology presented in this report for assessing model performance and hence comparing ‘simple’ and ‘advanced’ models in radioecology is developed based on the following aspects:

- Radioecological models need to be fit for purpose and a quality indicator can assert the model performance only with respect to its specific purpose.
- In radioecology, data often are unavailable, and in this case quality assessment can only be carried out qualitatively.
- If radioecological data are available, they are often characterised by large spatial and temporal variability.
- Data required for calculations of the considered metric need to be representative for the specific purpose of the model, i.e. they cannot be representative in general.

The principal numerical metric for scatter suggested in this report is the root mean squared logarithmic error (RMSLE) that can deal in a clever way with largely variable data. Other metrics, e.g. the bias, the geometric mean bias and the Pearson product moment correlation coefficient can help to obtain quantitative information regarding the capability of the model to avoid overestimation or underestimation of the data and to preserve data pattern. In addition to these metrics, a set of qualitative indicators is also presented, in order to be able to judge the quality of a model if data are lacking. Both qualitative and quantitative indicators are merged into a performance table to provide a practical structured tool for dealing with the assessment of model performance.

The reader should keep in mind that assessing model performance should be part of the iterative process of model development, and eventually the model should be modified based on the findings of the performance assessment.

4 DESCRIPTION OF THE MODELS

For predicting evolution or to anticipate potential effect of not-yet-implemented remediation actions models are necessary. Modelling the behaviour of the different radionuclides in any contaminated territory in combination with monitoring results when they exist can be improved conceptually by moving from empirical models to process-based models.

Several models are already available in the market, using different simplifications, as CROM, GRANIS, TREE4 or ERICA tool. These existing models have been tested against the data collected in the Territories Library Database (TLD). This report presents the results obtained with these existing models and more advanced models and their predictive power.

Under term “predictive power”, we indicate the correctness of values of any quantity predicted by the model. To quantify the predictive power, we have to follow these steps:

1. define a real-world scenario, which includes the necessary input quantities and corresponding output quantities for the model;
2. collect a data set of input values measured in the real world, which in the framework of Territories project is done under the TLD;
3. collect a data set of parameters, that corresponds to the model;
4. run the model with input parameters;
5. compare the model output with corresponding measurements in the scenario.

We can compare the performance of several models, when following the steps above with the same data sets. The comparison of models by means of predictive power is possible only with input and output quantities measured in the real world. Just an intercomparison between the models, based on any fictional scenario, is not sufficient to decide about their predictive power, even if good coherence between the models is observed, because all the models under consideration may be biased in a similar way – such a situation occurs very likely, if a certain underlying concept is not correct for defined scenario. The “fictional” intercomparisons, however, are not useless, as they reveal the inconsistencies between the models, initiating the revising of conceptions and algorithms. Fictional scenarios are not limited with measurements and thus, can be applied for a large number of different situations – this is their main advantage.

In practice, real-world scenarios in Territories are already existing contaminated sites and the measurements provided with main aim to make good decisions for reducing the harmful consequences. In some cases the source term (amount of radionuclides emitted into water or air, the site and time interval of emission) is known with sufficient precision. In such a case it is possible to start the model from the release and simulate all the pathways that lead to the measurable quantities, e.g. activity concentrations in water, soil, air, plants, animals and human food. The human exposure doses cannot be measured directly. The best proxy is based on groundshine, immersion, concentrations in food and drinking water. Thus, good knowledge on these quantities is a precondition for correct dose estimation and for a most useful validation of models as well.

The TLD datasets include activity concentrations measured in the accumulating media: soil and water bodies. If the measured concentrations in plants and animals are representative enough, then the biota intake modules of models can be tested.

4.1 EXISTING/WIDELY APPLIED MODELS

4.1.1 CROM

CROM incorporates the dispersion models included in the SRS19 (IAEA, 2001). The last version: CROM8 integrates the dose assessments for humans and biota, by incorporating the factors included in Erica tool (v 2015) for biota (see Main screen of CROM8 in Figure 4.1-1). The code in this version comprises parameters for 163 radionuclides for humans and 63 radionuclides for biota. CROM8 can therefore calculate activity concentrations in the environment using the IAEA recommended dispersion generic models and then estimate effective doses to humans and absorbed doses to biota, via different modules within the one code. The code also incorporates many default factors recommended for the calculation of the different exposure pathways for humans, as can be transfer factors to calculate the concentrations in foodstuff or dose conversion factors (derived from EPA FGR12, for ICRP 60 organs and SRS19 assumptions). These factors are widely used in other codes and are the more simple and basic models which can be applied to several aspects in radioecology.

In the case of biota calculations CROM also allows to use not only dispersion models, from the source term of one installation, but also environmental concentrations to assess the absorbed doses to biota. In this case the code incorporates two databases, one for the ERICA Ros and one for the ICRP RAPs.

New radioisotopes can be incorporated by the user, provided every parameter has been previously calculated.

A continuous QA of the CROM code is carried out by CIEMAT, in collaboration with several other organisations (RPD-HPA, IAEA, ARN, BfS and others). The code has been transferred to the IAEA in 2007 to be used as the reference code for generic assessments around installations for routine continuous discharges.



Figure 4.1-1: Main screen of CROM8.

Based in the same models, parameters and assumptions a new family, based in open software is being developed: the Open-Platform CROM (OPCROM). A first proof of concept based in OPCROM was the CROMERICA (<https://ciemat.cromerica.es>). However OP-CROM offers multiple possibilities, as including new mathematical models for additional not considered ecosystems or processes, needed for some specific assessments. New databases can be also easily included in the code, for new dose conversion factors, both for humans or biota or new factors associated to the models. All the parameters and factors associated to the code are in separate easily readable files. Also every module associated to a specific mathematical model for a process is coded as a separate text file. The same happens with the input and output files, being all of them separate easily readable files. Although multiple Graphical User Interfaces (GUI) can be developed to use OP-CROM (CROMERICA being one of them) a GUI is under development for an easy use of the code by non-experienced users and for a graphical representation of the results.

From the many dispersion models and aggregated factors incorporated in CROM, which are widely applied in many other models worldwide, only two aggregated factors could be applied within TERRITORIES project (CRs and DCFs). Additionally, based in the simplifications and assumptions used in the models incorporated in CROM, a new forest model (FORESTCROM) has been developed from scratch and tested in two sites included in the project. FORESTCROM is described below in a different chapter.

4.1.1.1 Use of CRs

For the comparison of the concentrations in the trees in two of the forests part of this project (Fukushima and the Belgian NORM site) two different approaches were followed. For the first one a new generic forest model has been developed and applied (FORESTCROM), and later explained. For the second one aggregated factors in equilibrium (transfer factors or concentration ratios) are traditionally used and incorporated in CROM. For the RAP pine tree, CR factors from Erica tool were incorporated in CROM and has been used to calculate activity concentration in the trees calculated from activity concentration in the soils. Transfer of radionuclides from soils to plants and animals has been traditionally described by aggregated factors (TFs and CRs) used to calculate radiation doses to them and to humans. In the case of CRs, the equation used is:

$$CR = \frac{C_{org}}{C_{soil}} \quad (4.1)$$

C_{org} is the activity concentration in any organism or its organ and C_{soil} is the activity concentration in the soil, measured in equilibrium conditions. ΔC_{org} and ΔC_{soil} are their uncertainties respectively.

Where the total uncertainty of the concentration ratio CR, is based on general formulation (Ku, 1966):

$$\Delta CR = \sqrt{\left(\frac{\Delta C_{org}}{C_{soil}}\right)^2 + \left(\frac{C_{org} \Delta C_{soil}}{C_{soil}^2}\right)^2} \quad (4.2)$$

ΔC_{org} and ΔC_{soil} can be based on standard deviations, if a number of soil and/or organism data of same type are available from a certain site; if concentrations in soil and same type of organisms are known from different sites, then it is possible to estimate the CR uncertainty between sites as the standard deviation of site-averaged CR-s.

The TLD data set for the Belgian NORM site includes activity concentrations of several radionuclides in soils and organisms (trees, grass plants, mosses, lichens and soil worms). To test the concentration ratio approach, the concentration ratios from CROM or ERICA tool were used.

4.1.1.2 Use of DCFs

Additionally the data from Fen site were used to compare ambient dose equivalent rates obtained using simple dose conversion factors (DCFs) implemented in the CROM code (calculated from the unit conversion factors derived from the EPA FGR12 report (EPA, 1993), derived by CIEMAT for the ICRP 60 organs, using the IAEA SRS 19 assumptions), which convert radioactive deposit in a soil ($Bq\ m^{-2}$) to ambient dose equivalent ($Sv\ y^{-1}$), with directly measured ambient dose equivalent *in situ*.

4.1.2 NORMALYSA

NORMALYSA (NORM And LegacY Site Assessment) has been described previously by Avila *et al.* (2018) and comprises of a set of models and databases covering:

- Relevant sources for NORM and legacy sites
- Transport pathways (airborne, groundwater etc.)
- Contaminated receptor environments

NORMALYSA is comprised of relatively simple transport/ exposure models based upon SRS-19 (IAEA, 2001) with their intended application for routine releases of radioactivity (often most amenable to small facilities). The databases used within the programme are from reputable data collations, e.g. IAEA TRS-472 (IAEA, 2010). A schematic of the ‘receptor’ model for forests is shown in Figure 4.1-2.

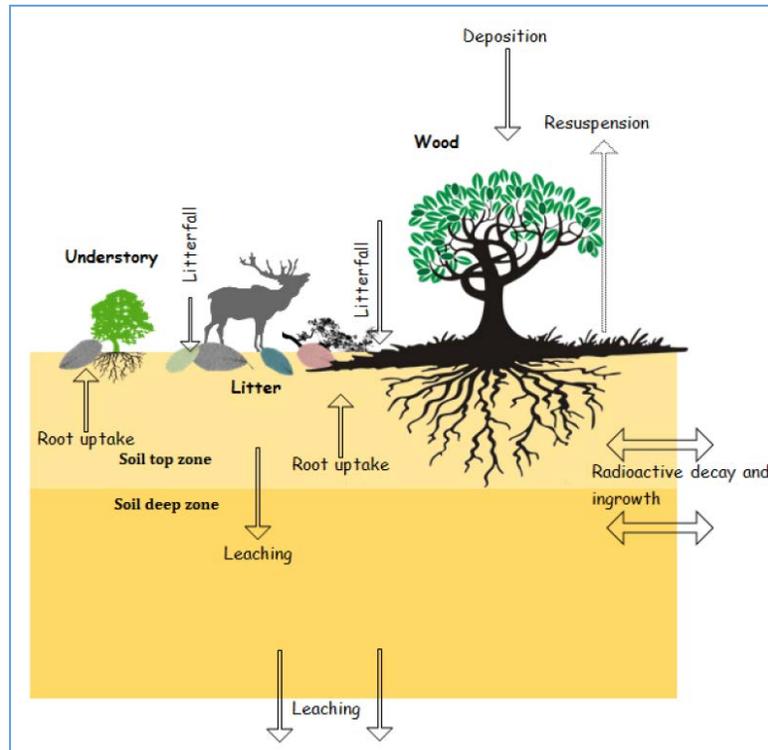


Figure 4.1-2: Illustrative scheme of the ‘Forest’ module. Exchanges between media, loadings and losses are shown by arrows. Reproduced from Avila *et al.* (2018).

Input information needed is typically in the form of activity released from a source per unit time or starting with radionuclide activity concentrations in waste or contaminated soil. Numerous default data underlie the various models used in NORMALYSA and further details can be found in Avila *et al.* (2018).

The model, for terrestrial systems, is primarily configured to simulate an atmospheric source and input. For example, one can define a stack release and then use, for example, a model such as SRS-19 Gaussian plume model (IAEA, 2001) to simulate advection and turbulent diffusion of a given radionuclide. This type of model was not appropriate for our particular case (i.e. with the intended application to Fen, Norway) because we are not dealing with a system characterised by a source followed by radionuclide transport to a distant location. The information we have available (from the TLD) relates specifically to radionuclide concentration levels currently present in the environment. The atmospheric input of radionuclides is putatively minor (albeit with the knowledge that some input of ^{210}Pb and ^{210}Po at natural levels will be occurring) compared to the levels that already exist in soils with their provenance attributable to the underlying, highly mineralised rock formations in the area where the model has been applied. NORMALYSA actually has a system in place to enter measured data directly – under assessment “options”. This allowed radionuclide activity concentration data for soils, from the TLD, to be entered specifically.

As noted earlier in this report (Section 3.1), interaction matrices are often employed when developing conceptual models for a given case. NORMALYSA allows for this form of expressing the model structure (Figure 4.1-3) allowing the main processes to be explicitly shown and the limitations (cf. a broad consideration of potential pathways for forest ecosystems as presented from a purely conceptual perspective – Section 3.1) to be expressed.

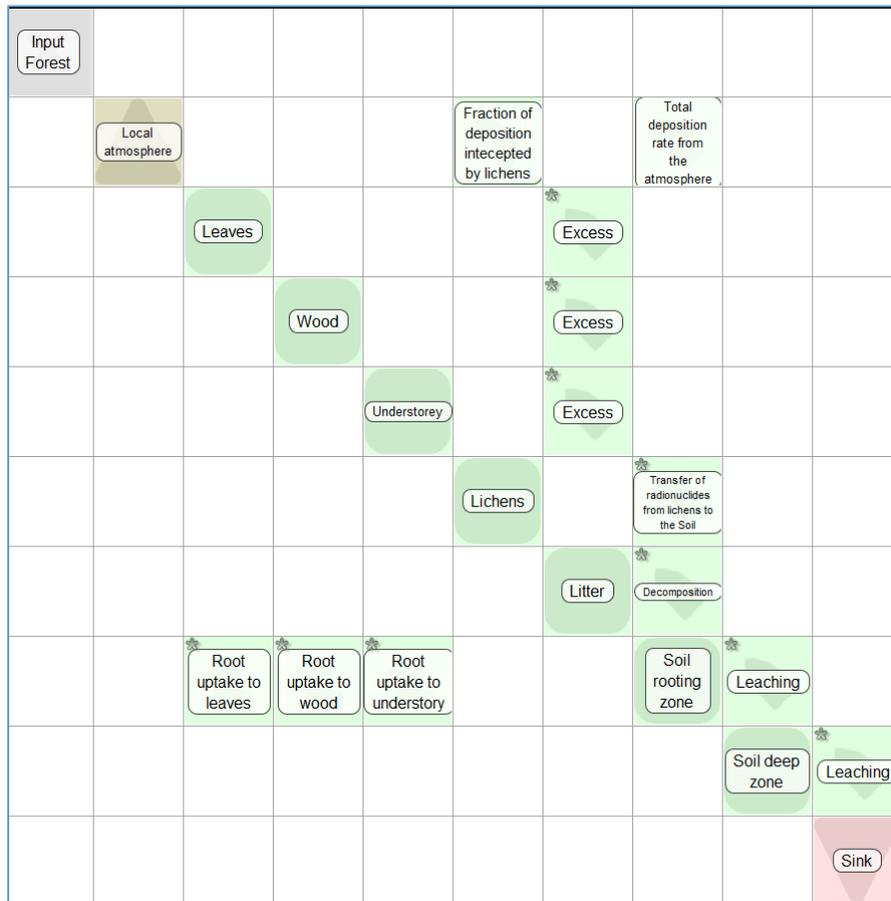


Figure 4.1-3: Forest receptor model from NORMALYSA showing the main processes considered in the model.

NORMALYSA has four categories that are relevant in case of using plant and soil data from the TLD: wood, leaves, understorey and berries. Different plant samples from Fen have been mapped onto the plant categories in NORMALYSA in the following way:

- Tree (represented by two categories: wood and leaves) → spruce, pine, birch,
- Understorey → moss, lichen, fern, dandelion, grass
- Berries → strawberry

Radionuclides transfer from soil to the above-mentioned environmental components is considered using a concentration ratio approach [e.g. as described in IAEA, 2014].

A mass balance approach has been adopted to address the dynamics of the system following a deposition event, with radionuclide fluxes modelled as first order ‘transfer’ process including fluxes:

- from top soil to tree wood via root uptake,
- from top soil to tree leaves via root uptake,
- from top soil to understorey (plants and mushrooms) via root up take,
- from tree leaves to litter by leaves fall,
- from tree wood to litter by wood fall,
- from understorey plants to litter by plant senescence,
- from litter to top soil layer by following litter decomposition.

Radionuclide concentrations in game animals are calculated using a transfer coefficient approach based on radionuclide concentrations in mushrooms, berries, leaves and wood allowing for the animal’s diet.

The mathematical representation of this entire system would be laborious and an exhaustive presentation is considered unnecessary here. A comprehensive overview can be found in Avila *et al.* (2018). However, by way of example, the mass balance for radionuclides in the forest canopy (tree leaves) is given by the differential equation (4.3):

$$\frac{dA_{leaf}}{dt} = A_{soilRZ} \times TC_{root>leaf} - A_{leaf} \times TC_{leaf>litter} - \lambda \times A_{leaf} + \sum_{p \in P_i} Br_p \times \lambda \times A_{leaf} \quad (4.3)$$

Where:

A_{Leaf} = radionuclide inventory in leaves including yearly and older leaves (Bq),

A_{soilRZ} = radionuclide inventory in the soil root zone (Bq),

$TC_{root>leaf}$ = Transfer coefficient from soil to leaves due to root uptake (1/year)

$TC_{leaf>litter}$ = Transfer coefficient from leaves to litter (1/year)

The two last terms on the right hand side of the equation (those including λ and Br_p) describe radioactive decay and ingrowth of radionuclides from parent nuclide(s).

Similar expressions are used to define fluxes between the various compartments in the model mentioned above (and as shown Figure 4.1-3).

The other pertinent calculation to note is the derivation of effective dose rate outdoors from deposited radionuclides or radionuclides present in waste material. The effective dose rate outdoors from the deposited radionuclides (\dot{D}_{eff} , Sv/h) is calculated by:

$$\dot{D}_{eff} = \sum_1^n \dot{D}_{eff,RN} \quad (4.4)$$

Whereby the dose rate for external irradiation from each radionuclide ($\dot{D}_{eff,RN}$, Sv/h) is calculated with the equation:

$$\dot{D}_{eff,RN} = C_{soil,RN} \times DCC_{ext,RN} \quad (4.5)$$

Where:

$C_{soil,RN}$ = radionuclide volumetric concentration in soil (provided by the respective receptor module) (Bq/m³),

$DCC_{ext,RN}$ = is the age group and radionuclide specific dose conversion factor for external irradiation ((Sv•m)/(Bq•h)).

The radionuclide specific dose conversion factors, $DCC_{ext,RN}$, have been taken from EPA (1993). Furthermore, a correction factor accounting for dose rate attenuation depending on thickness and density of cover material can be applied as described in Avila *et al.* (2018).

The NORMALYSA model allows for data to be entered in the form of probability distributions and for subsequent calculations to be run probabilistically. Using a Monte Carlo code, the model can be run based on one random value from each probabilistic input (i.e. where a distribution has been defined), producing one corresponding estimate of the model outputs. The process is then typically repeated hundreds or thousands of times to create a statistical sample of model outputs (i.e. the output is also produced in the form of a probability distribution).

4.1.3 FORESTCROM

CROM (Mora *et al.*, 2015) follows SRS 19 dispersion models (IAEA, 2001), which are generic models originally designed to obtain conservative results of effective doses to humans from routine discharges in planned situation, including the contamination and consumption of agricultural products.

These models cannot be explicitly applied to the case of an accidental situation (e.g. Fukushima accident) in the short term or to an existing situation (e.g. Tessengerlo NORM site) for the very long term. They cannot also be applied directly to forests due to the evident differences with agricultural products. However, the philosophy of developing simple, conservative approaches for these cases is still possible.

In the case of trees, which are an important part of a forest, we can expand the approach of applying a transfer factor - TF (or more properly a concentration ratio - CR) from soil to tree, which can be seen as the simpler model. TFs or CRs are defined as the ratios when the equilibrium is reached in the system. However, equilibrium is not achieved in the short term in any of the situations mentioned above.

A more complex model should consider the time dependence of the different processes contributing to the contamination in the tree, even if considered as a whole (see Figure 4.1-4). Although in the figure many other processes are shown, in a second approach not only the root absorption, but also the deposition from the atmosphere on the trees and on the soil were considered. Time dependence of all of those processes was considered. Moreover, every parameter included in the model (e.g. deposition velocity or growth of trees) can be time dependent. Taking this into account, a new model named FOREST-CROM was derived.

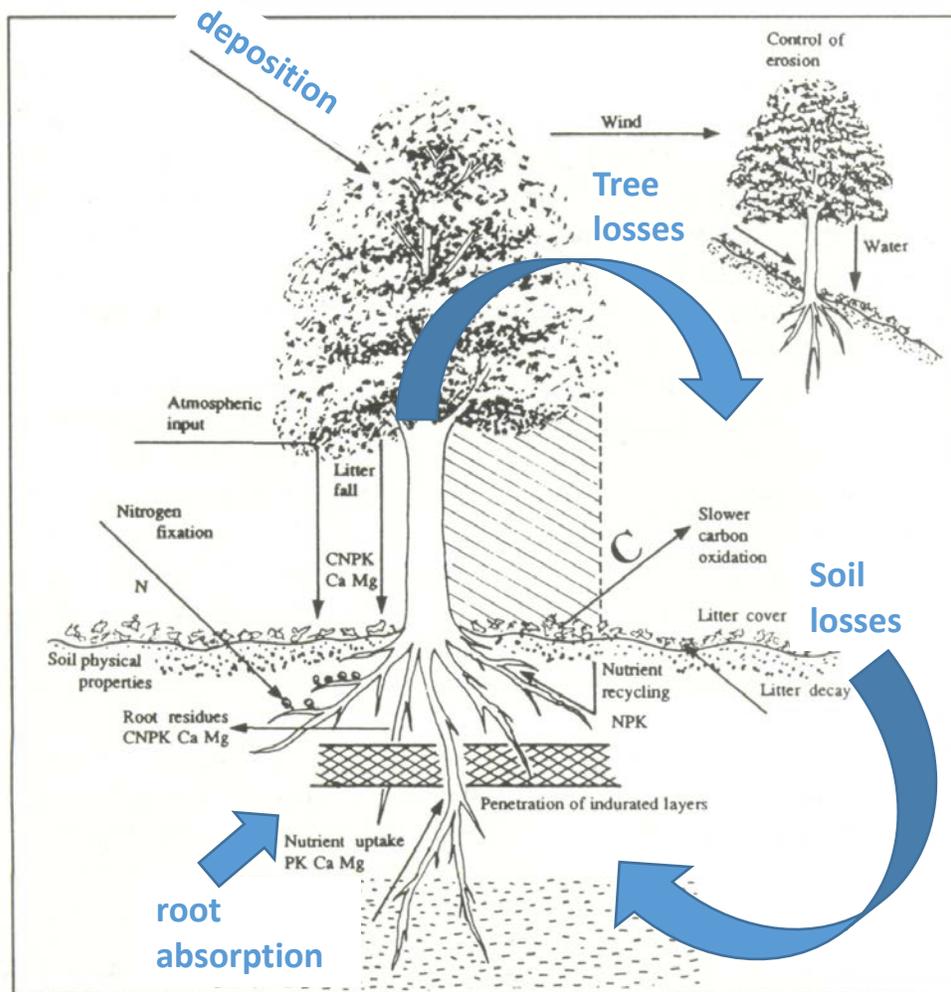


Figure 4.1-4: Simplified view of the behavior of chemical elements (or radionuclides) in the soil-tree system.

For deriving a differential equations describing the system, we can simplify it by considering (see Figure 4.1-5): (1) the soil as a single, homogeneously-contaminated layer; (2) the outer part of the tree (external contamination of the tree); and (3) the inner part of the tree (wood, fruits, leaves or needles,...), in this case considered homogeneously contaminated.

There is a possibility, in the studied cases, of an initial outer contamination of the tree, and here is when this outer part of the tree becomes important. However, it is also an important case when a new tree grows in contaminated soil (either with artificial or natural radioisotopes). Both situations are included in the model, considered the initial conditions.

The total activity of the tree ($C_{tree}(t)$) can be divided into the inner part of the tree ($C_{tree-inn}(t)$) and the external contamination of the tree ($C_{tree-out}(t)$), which can be lost by different phenomena (e.g. effects of rain, snow or wind), as will be discussed below.

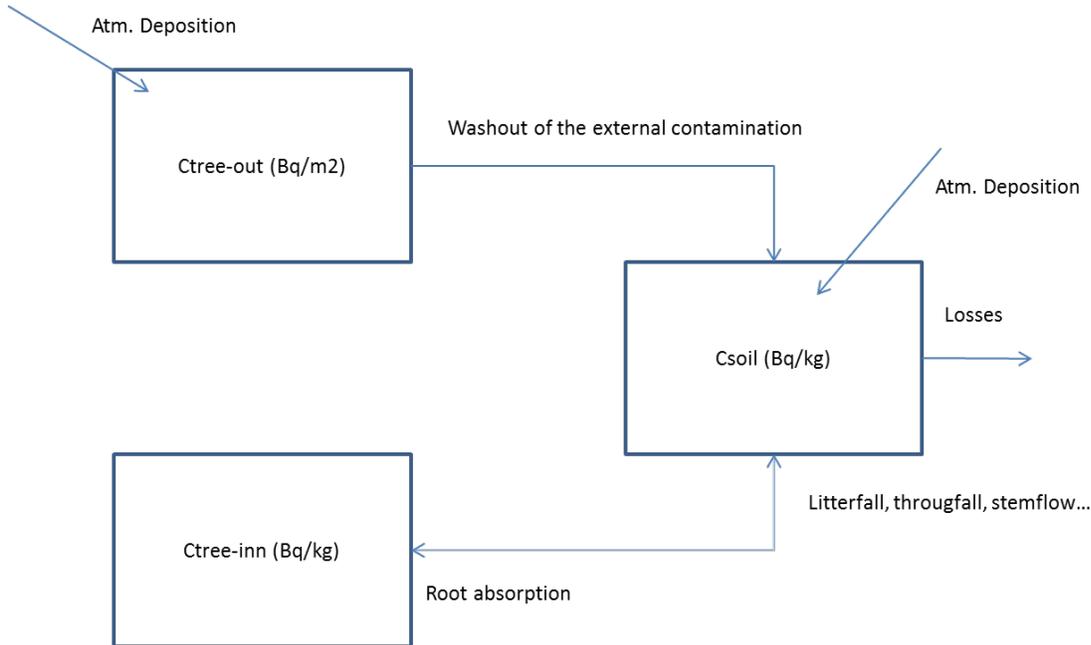


Figure 4.1-5: Conceptual model of the simplifications assumed.

In view of those simplifications, the differential equation which can be derived for obtaining the internal activity concentration in a tree, absorbed from the roots from a known homogeneous soil activity concentration, is:

$$\frac{dC_{tree-inn}}{dt} = k_{s-t} \cdot C_{soil} - \lambda_{tree} \cdot C_{tree-inn} \quad (4.6)$$

Where:

$C_{tree-inn}$ is the concentration in the inner part of the tree for a given time t (Bq/kg), considering tree as a whole organism;

C_{soil} is the concentration in the soil for a given time t (Bq/kg)

K_{s-t} is a transfer coefficient from soil to tree, considered constant (d^{-1})

$\lambda_{tree} = \lambda_{phy} + \lambda_{ec-t}$ is the sum of radionuclide decay constant (λ_{phy}) and a ecological decay constant (λ_{ec-t}), dependent on the type of tree, soil, chemical element or chemical compound, and other conditions. Therefore λ_{tree} is an ecological decay constant depending on the type of tree and isotope (d^{-1}).

In this case $C_{tree-inn}$ includes all the parts of the organism: wood, fruits, leaves (or needles), bark, sap, etc. which are considered homogeneously contaminated.

This means that there is a constant rate of transfer from the soil to the inner part of the tree, due mainly to the root absorption of nutrients (for instance, K, which has an analogous metabolic behavior to Cs), is considered an instantaneous distribution to all the different parts of the tree and a constant rate of loss of the activity concentration in the tree due to several phenomena (for instance the falling of fruits, leaves, etc.).

To solve the differential equation (4.6), k_{s-t} , λ_{tree} and $C_{soil}(t)$ should be derived, measured or modeled.

For deriving the k_{s-t} we can apply the definition of TF or CR, i.e. ratios in equilibrium conditions ($CR = \frac{C_{tree-inn}}{C_{soil}}$). By definition, TF or CR, (unitless) are measured in equilibrium and can be obtained in the bibliography (IAEA, 2010; IAEA, 2014). Moreover, in equilibrium $\frac{dC_{tree-inn}}{dt} = 0$, therefore, from eq. (4.6) $k_{s-t} \cdot C_{soil} = \lambda_{tree} \cdot C_{tree-inn}$. And so, as $CR = \frac{C_{tree-inn}}{C_{soil}} \rightarrow CR = \frac{\lambda_{tree}}{k_{s-t}} \rightarrow [k_{s-t} = \lambda_{tree} \cdot CR]$, for given conditions of a given tree specie, a given soil and a given radionuclide. This can be substituted in eq. (4.6):

$$\frac{dC_{tree-inn}}{dt} = \lambda_{tree} \cdot [CR \cdot C_{soil} - C_{tree-inn}] \quad (4.7)$$

λ_{tree} , as was proposed for agricultural products in IAEA SRS 19 (IAEA, 2001), can be divided in two components, one physical, only dependent on the radionuclide decay constant (λ_{phy}) and one ecological, dependent on the type of tree, soil, chemical element or chemical compound, and other conditions: (λ_{ec}). Therefore $\rightarrow \lambda_{tree} = \lambda_{phy} + \lambda_{ec-t}$. For the first, the radioactive decay constant: λ_{phy} (d⁻¹) is a well-documented value which can be found in several databases. Values for the second one, λ_{ec-t} , also can be found, as those based for instance in Chernobyl accident, and can be applied to test the model in other situations as Fukushima or Tessengerlo.

For the external contamination of the tree another differential equation describes the variations:

$$\frac{dC_{tree-out}}{dt} = \alpha \cdot \dot{d} - k_{t-s}^{out} \cdot C_{tree-out} \quad (4.8)$$

Where:

$C_{tree-out}$ is the concentration in the external part of the tree for a given time t (Bq/m²);

α is an interception factor, which considers the part of the deposition which is intercepted by the external part of the tree (unitless);

\dot{d} is the deposition rate (Bq/m²/d), which can also be a function of time;

k_{t-s}^{out} is a constant transfer coefficient from the external part of the tree to the soil (d⁻¹) – due to wash-out of the external contamination for instance;

α is an interception factor (or canopy interception fraction), considering the part of the deposition which is intercepted by the external part of the tree (unitless);

For the external contamination of the tree several possibilities can be considered. For instance a continuous and constant deposition rate could simulate routine discharges from a nuclear facility, an

instantaneous total deposition at a given time (e.g. $t=0$) and a deposition rate null at any other time (a delta function), or any others $\dot{d}(t)$.

In that last case, equivalent to an accident, the solution is simplified to:

$$C_{tree-out}(t) = C_{tree-out}(0) \cdot e^{-k_{t-s}^{out} \cdot t} = \alpha \cdot d(0) \cdot e^{-k_{t-s}^{out} \cdot t} \quad (4.9)$$

Where the initial deposition $d(0)$, which depends on the location, should be obtained either by measurements carried out in $t = 0$; either by back-calculations based on time series of measurements; or modelled from available data of the accident together with appropriate dispersion models as JRODOS. The interception factor α depends on the density and age of the forest (number of trees per Ha), the type of tree and its density of leaves, and other factors. However, this was well documented for coniferous forest after Chernobyl accident (IAEA, 2010), being the range for pine forests from 0.4 to 1 (average 0.7). The last factor needed, k_{t-s}^{out} , depends on many effects, such as the rain rate at a given location, but can be considered as a constant factor averaged over time (a constant loss of the external contamination).

Finally, it is necessary to describe the activity concentration in the soil ($C_{soil}(t)$). In reality, soil can show variations in the activity concentrations at different depths due to many effects as evaporation, deposition in the surface, leaching, etc. Usually all these effects are modeled as different compartments representing different layers of depth. As a first approach we assume the concentration in the soil being homogeneous in the depth affecting the transfer to the trees (i.e. roots layer considered as a 1-layer approach). Note that other models (like ECOFOR model) use a "several layers" approach. In this single homogeneously contaminated layer case, the variation of the concentration with time would be:

$$\frac{dC_{soil}}{dt} = \frac{1}{\rho \cdot h} \cdot [(1 - \alpha) \cdot \dot{d} + k_{t-s}^{out} \cdot C_{tree-out}] + \lambda_{tree} \cdot C_{tree-inn} - \lambda_{soil} \cdot C_{soil} - \lambda_{tree} \cdot CR \cdot C_{soil} \quad (4.10)$$

(16)

Where:

$C_{tree-inn}$ is the concentration in the inner part of the tree for a given time t (Bq/kg);

$C_{tree-out}$ is the concentration in the external part of the tree for a given time t (Bq/m²);

C_{soil} is the concentration in the soil for a given time t (Bq/kg);

α is an interception factor, considering the part of the deposition which is intercepted by the external part of the tree (unitless);

ρ is the density of the soil (kg/m³);

h is the depth of the soil (m);

\dot{d} is the deposition rate (Bq/m²/d);

λ_{tree} is a decay constant (radioactive+ecological), where the ecological decay coefficient from inner part of the tree (bark, wood, fruits, leaves) to the soil (d^{-1}) is due to several effects, as litterfall, throughfall, or stemflow;

k_{t-s}^{out} is a constant transfer coefficient from the external part of the tree to the soil (d^{-1}) – due to wash-out of the external contamination for instance;

λ_{soil} is an ecological decay constant depending on the type of soil and isotope (d^{-1}).

For the general case all of the parameters should be determined in advance, although very limited information is available and only for few radionuclides. However there are some cases where additional simplifications can be used.

The simplest case to solve this equation happens when $\lambda_{tree} \cdot C_{tree-inn} + \frac{k_{t-s}^{out}}{\rho \cdot h} \cdot C_{tree-out} \cong 0$, meaning that the contribution to the soil activity concentration is negligible for any activity falling from the tree, by any process, and the external deposition rate $\dot{d} = 0$. In that case, the equation becomes:

$$\frac{dC_{soil}}{dt} = -\lambda_{soil} \cdot C_{soil} \quad (4.11)$$

Having an immediate analytical solution where:

$$C_{soil}(t) = C_{soil}(0) \cdot e^{-\lambda_{soil} \cdot t} \quad (4.12)$$

Solving the system of equations above can be done numerically, for any variation with time of any of the parameters (even for constant parameters) and for given initial conditions.

For λ_{soil} there are also several options. For the case of long half-life radioisotopes, as some of the radionuclides in the natural decay chains, in secular equilibrium and in a pond, as is the case in Tessengerlo, a constant concentration in the soil can be assumed, with no external contamination of the trees or any other external deposition rate. However, for artificial radionuclides originating in accidents (as Cs^{137} after Fukushima accident) a loss must be included

In the IAEA SRS 19 (IAEA, 2001) for instance, a recommendation is provided for $\lambda_{soil} = \lambda_{phy} + \lambda_{ec-s}$, where the λ_{phy} is the radioactive decay constant of a given radionuclide, and the λ_{ec-s} takes into account several phenomena as wash-out of the deposited material in the surface of the soil, leaching and other effects.

SIMPLIFICATIONS OF THE MODEL

- Single layer homogeneously contaminated soil;
- Homogeneous characteristics of the soil (density or contaminated depth affecting the roots);
- use of CR, factor aggregating many incompletely understood processes; Homogeneously contaminated tree, considered as a whole organism;
- Two parts of the contamination in the tree are considered independently: internal and external contamination, with different units;
- Use of ecological decay factors λ_{ec} , in trees and in soils, aggregating many incompletely understood processes and with limited data for only some radionuclides;

Initial concentrations in each component of the model: initial C_{soil} is inferred in the case of Fukushima, initial $C_{\text{tree-inn}}$ in the case of the natural radionuclides cannot be 0, as there would always be a natural background from any other type of soil, this can be improved by the use of the average natural concentration of the natural decay chains in the soils – around 30 Bq/kg worldwide average for U-238, according with UNSCEAR reports.

UNCERTAINTIES OF THE MODEL

Several aggregated factors like CRs or ecological half-lives are provided with parameters to develop a probability density function (pdf), which can be used to perform appropriate uncertainty propagation through the model using Monte-Carlo methods for instance. The conceptual model uncertainty (Urso *et al.*, 2019) must be also evaluated.

4.1.4 TREE4-simple

The TREE4-simple forest transfer model implemented in the IRSN's SYMBIOSE platform was also used as simple model for the Fukushima sites. This model was originally developed after Chernobyl in the course of the RODOS project (RODOS, 1999; Ehrhardt and Weis, 2000; Rantavaara *et al.*, 2001) and further tested against other European approaches in the course of the IAEA BIOMASS & EMRAS programs (Shaw *et al.*, 2005; Calmon *et al.*, 2009). In the first few years after Fukushima accident, this model was slightly improved and tested against the first few published data on radiocesium contamination in Japanese forests (Calmon, 2014; Calmon *et al.*, 2015; Gonze *et al.*, 2016).

As depicted in Figure 4.1-6, this model consists of a single module dedicated to the prediction of the recycling of radionuclides (^{137}Cs , here) within a soil-tree subsystem contaminated by atmospheric fallouts. Transfers to understorey vegetation, berries, mushrooms and wild game are not investigated. The Cs module is responsible for calculating the daily evolution of ^{137}Cs concentrations (Bq kg^{-1} dry mass), inventories (Bq m^{-2}) and fluxes ($\text{Bq m}^{-2} \text{d}^{-1}$) in this medium, from days to decades after an initial atmospheric deposit. The calculations rely on the knowledge of the deposition characteristics (e.g. deposition rate, dry/wet deposition ratio, rainfall height) as well as the forest stand characteristics (e.g. biomasses, area indexes, biomass renewal rate). In this simple approach, the ageing of the forest stand is neglected, i.e. the stand characteristics are assumed constant with time.

The conceptual model describing the compartments and transfer processes which are accounted for in the Cs module is depicted in Figure 4.1-7. In this simple approach, the foliage and branch compartments are not distinguished and aggregated into a single "canopy" compartment. This simplification prevents us from calculating activity concentrations separately and comparing them with observations. The canopy compartment is splitted into an "external" pool, directly impacted by the atmospheric fallouts, and an "internal" pool contaminated by foliar incorporation and root uptake. The forest floor is decomposed into an upper "organic" layer (e.g. litter OI and fragmented Of horizons) and an underlying "mineral" soil (e.g. humified Oh and mineral horizons). We assume that the "mineral" soil consists of a single homogeneous layer, which prevents us from simulating the depth distribution of ^{137}Cs in the soil profile and comparing it with observations. Other simplifications exist which we will discuss later, in the light of the advanced TREE4 model (see section 7.2.2).

The mathematical model relies on the resolution of a system of non-stationary differential equations (i.e. inventory-flux equations) which expresses (^{137}Cs) mass conservation in the system. Concentration of ^{137}Cs in each compartment is deduced from the predicted inventory by dividing it by the corresponding (bio)mass ($\text{kg dry mass m}^{-2}$) which therefore must be known. The parameterisation of the transfer processes relies on the use of constant kinetic rates (in day^{-1}), with default generic values mostly derived from post-Chernobyl observations.

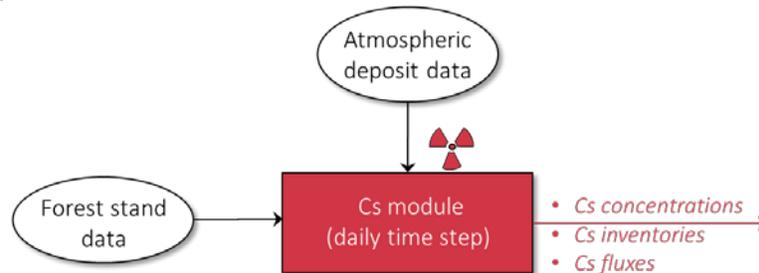


Figure 4.1-6: TREE4-simple model structure (implemented in the IRSN's SYMBIOSE platform). See main text for description.

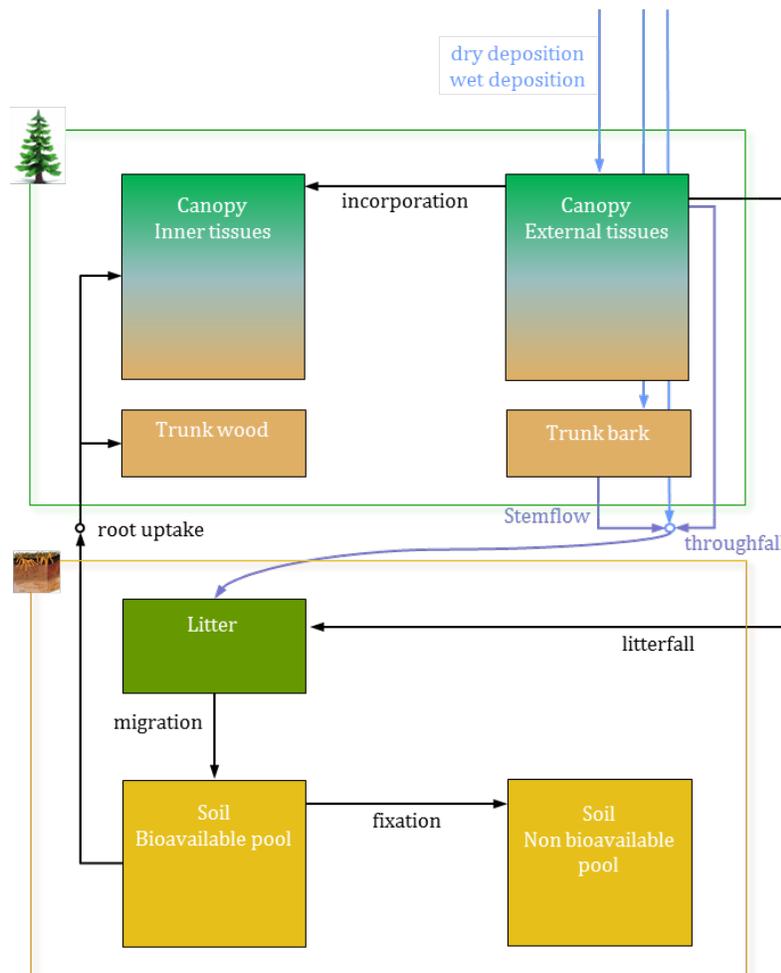


Figure 4.1-7: Conceptual modelling of Cs cycling within the soil-tree system (TREE4-simple). See main text for description

4.1.5 NRPA box modelling approach (ARCTICMAR)

The present model uses a modified approach for compartmental modelling (Iosjpe *et al.*, 2002, 2009; Iosjpe, 2006) which allows the study of dispersion of radionuclides over time (non-instantaneous mixing in the oceanic space). The box structures for surface, mid-depth and deep water layers have been developed based on the description of polar, Atlantic and deep waters in the Arctic Ocean and the Northern Seas and site-specific information for the boxes generated from the 3D hydrodynamic model NAOSIM (Karcher and Harms, 2000). Surface structure of the ARCTICMAR model is presented in Figure 4.1-8.

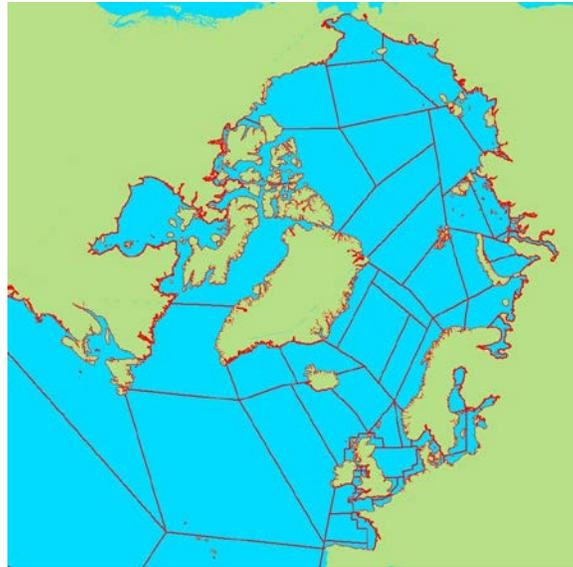


Figure 4.1-8: The structure of the surface water boxes for the NRPA box models (ARCTICMAR -345 compartments).

The box model includes the processes of advection of radioactivity between compartments, sedimentation, diffusion of radioactivity through pore water in sediments, particle mixing, pore water mixing and a burial process of radioactivity in deep sediment layers. Radioactive decay is calculated for all compartments. Accumulation of contamination by biota is further calculated from radionuclide concentrations in filtered seawater in different water regions. Doses to humans are calculated on the basis from given seafood consumptions, based on available data for seafood catches and assumptions about human diet in the respective areas. Dose rates to biota are developed on the basis of calculated radionuclide concentrations in marine organisms, water and sediment, using dose conversion factors. Its structure is presented in Figure 4.1-9.

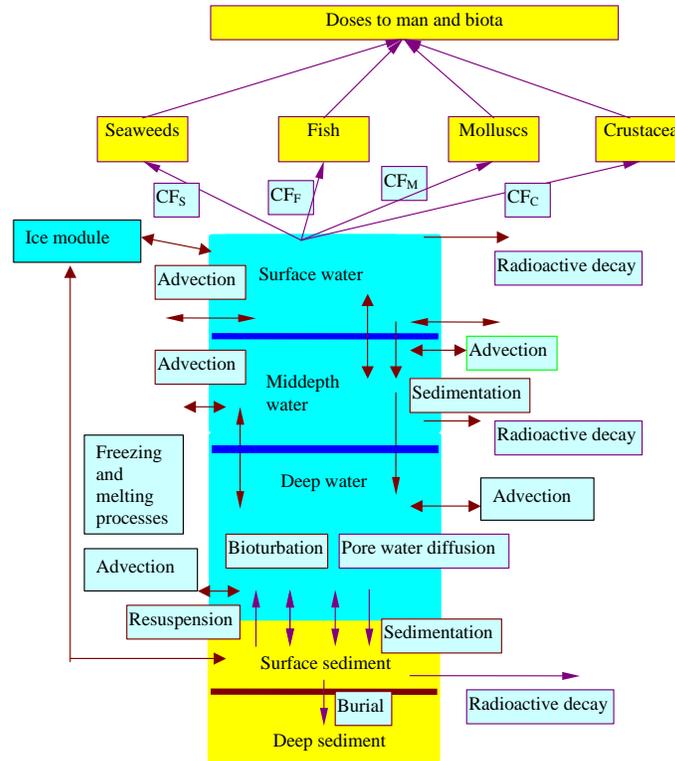


Figure 4.1-9: A schematic structure of the processes involved in modelling.

The equations of the transfer of radionuclides between the boxes are of the form:

$$\frac{dA_i}{dt} = \sum_{j=1}^n k_{ji} A_j \gamma[t \geq (T_j + w_{ji})] - \sum_{j=1}^n k_{ij} A_i \gamma[t \geq (T_i + w_{ij})] - k_i A_i \gamma(t \geq T_i) + Q_i, t \geq T_i \quad (4.13)$$

$$A_i = 0, t < T_i$$

where $k_{ii}=0$ for all i , A_i and A_j are activities (Bq) at time t in boxes i and j ; k_{ij} and k_{ji} are rates of transfer (γ^{-1}) between boxes i and j ; k_i is an effective rate of transfer of activity (γ^{-1}) from box i taking into account loss of material from the compartment without transfer to another, for example radioactive decay; Q_i is a source of input into box i ($\text{Bq } \gamma^{-1}$); n is the number of boxes in the system, T_i is the time of availability for box i (the first times when box i is open for dispersion of radionuclides) and γ is an unit function:

$$\gamma(t \geq T_i) = \begin{cases} 1, & t \geq T_i \\ 0, & t < T_i \end{cases} \quad (4.14)$$

The times of availability T_i

$$T_i = \min_{\mu_m(v_0, v_i) \in M_i} \sum_{j,k} w_{jk} \quad (4.15)$$

are calculated as a minimized sum of the weights for all paths $\mu_m(v_0, \dots, v_i)$ from the initial box (v_0) with discharge of radionuclides to the box i on the oriented graph $G=(V, E)$ with a set V of nodes v_j correspondent to boxes and a set E of arcs e_{jk} correspondent to the transfer possibility between the boxes j and k (graph elements as well as available paths are illustrated by Figure 4.1-10). Every arc e_{jk} has a weight w_{jk} which is defined as the time required before the transfer of radionuclides from box j to box k can begin (without any way through other boxes). Weight, w_{jk} , is considered as a discrete function F of the water fluxes f_{jk} , f_{kj} between boxes j and k , geographical information g_{jk} and expert evaluation X_{jk} . M_i is a set of feasible paths from the initial box (v_0) to the box i (v_i).

The traditional box modelling is a particular case of the present approach when all times of availability in (1) are zero: $\{T_i\} = 0, i = 1, \dots, n$.

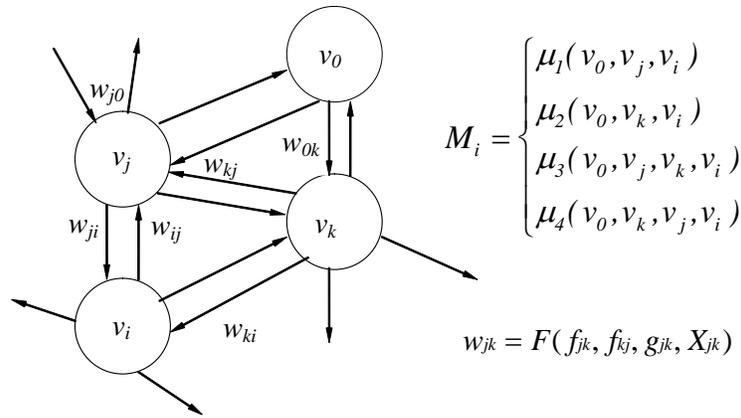


Figure 4.1-10: Graph elements.

Expressions for the transfer rates of radioactivity between the bottom water and sediment compartments will be useful in the present analysis (the transfer rates are shown in Figure 4.1-11):

$$k_{WS} = \frac{SR \cdot k_d}{d \cdot (1 + k_d \cdot SSL)} + \frac{D}{d \cdot h_S (1 + k_d \cdot SSL)} + \frac{R_T \cdot \omega \cdot h_S}{d \cdot (1 + k_d \cdot SSL)} + \frac{R_W \cdot \rho \cdot k_d \cdot (1 - \omega)}{d \cdot (1 + k_d \cdot SSL)} \quad (4.16)$$

$$k_{SW} = \frac{D}{h_S^2 \cdot [\omega + k_d \cdot \rho \cdot (1 - \omega)]} + \frac{R_T \cdot \omega}{\omega + k_d \cdot \rho \cdot (1 - \omega)} + \frac{R_W \cdot \rho \cdot k_d \cdot (1 - \omega)}{h_S \cdot [\omega + k_d \cdot \rho \cdot (1 - \omega)]} \quad (4.17)$$

$$k_{SM} = \frac{D \cdot \omega}{h_S^2 \cdot [\omega + k_d \cdot \rho \cdot (1 - \omega)]} + \frac{k_d \cdot SR}{h_S \cdot [\omega + k_d \cdot \rho \cdot (1 - \omega)]} \quad (4.18)$$

$$k_{MS} = \frac{D \cdot \omega}{h_S h_{SM} \cdot [\omega + k_d \cdot \rho \cdot (1 - \omega)]} \quad (4.19)$$

$$k_{MD} = \frac{k_d \cdot SR}{h_{SM} \cdot [\omega + k_d \cdot \rho \cdot (1 - \omega)]} \quad (4.20)$$

Here k_{WS} is composed of expressions describing the transfer of activity by sedimentation, molecular diffusion, pore water mixing and particle mixing, respectively. Similarly, k_{SW} is composed of expressions describing the transfer of radioactivity by molecular diffusion, pore water mixing and particle mixing. k_{SM} is composed of expressions describing the transfer of radioactivity by sedimentation and molecular diffusion. k_{MS} corresponds to the transfer by molecular diffusion. Finally, k_{MD} corresponds to the transfer of radioactivity by sedimentation. R_W ($m\ y^{-1}$) is the sediment reworking rate; R_T (y^{-1}) is the pore-water turnover rate; k_d ($m^3\ t^{-1}$) is the sediment distribution coefficient; SSL ($t\ m^{-3}$) is the suspended sediment load in the water column; SR ($t\ m^{-2}\ y^{-1}$) is the sedimentation rate; D ($m^2\ y^{-1}$) is the molecular diffusion coefficient, h_s (m) and h_{SM} (m) are the surface and middle sediment thickness respectively; ω is the porosity of the bottom sediment; ρ ($t\ m^{-3}$) is the density of the sediment material and d is the depth of the water column (see Table 4.1-1 and 4.1-2).

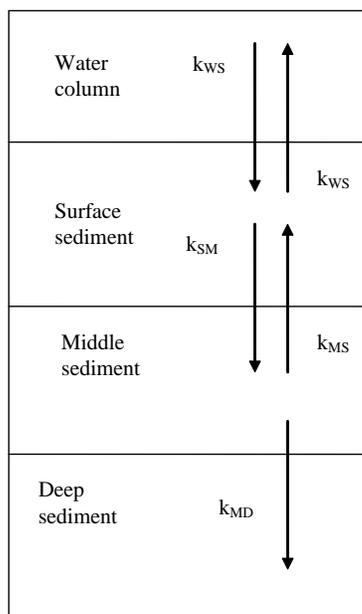


Figure 4.1-11: Generic vertical structure of the water-sediment compartments.

Table 4.1-1: Factors used in the NRPA Box model for Cs-137.

	Kdco	Kdoc	Kdap	CFfi	CFmo	CFcr	CFsw	Dfig	DFih	
	3.09e+1	4e+3	2e+3	230	1e+2	5e+1	6e+1	5e+1	1.3e-08	8.5e-09

HL is Half-Life of Cs-137; Kdco, Kdoc and Kdap - sediment distribution coefficients for the coastal waters, open ocean and apparent value for regions near sources of releases, correspondently; CFfi, CFmo, CFcr and CFsw – concentration factors fish, mollusks, crustaceans and seaweeds, correspondently; Dfig and DFih – dose conversion factors for ingestion and inhalation, correspondently.

Table 4.1-2: Additional environmental parameters

h=0.1 m	The thickness of the surface sediment layer
H=1.9 m	The thickness of the mid-depth sediment layer
Rw=5E-3 m/y	The sediment reworking rate
Rt= 1 y⁻¹	The pore-water turnover rate
ω=0.7	Porosity of bottom sediment
ρ=2.5 t/m³	Density of the bottom sediment

4.2 MORE ADVANCED MODELS

4.2.1 GRANIS

The computer model GRANIS (Gamma Radiation Above Nuclides In Soil) (Kowe, Carey *et al.* 2007) can calculate gamma dose and dose rate above layers of contaminated material of finite thickness and infinite lateral extent. By considering external irradiation from layers of contaminated material, the model allows variations in radionuclide activity concentrations with depth to be considered. For the TERRITORIES project, the model has been used to assess the dose rate above land contaminated with NORM. For this type of assessment measurements of radionuclide activity concentrations have been entered directly into GRANIS to enable the calculation of an instantaneous dose rate. The model includes the processes of radionuclide decay and progeny ingrowth. Inter-comparisons have been made of GRANIS results with those of standard theoretical methods, published papers and other radiation transport codes. These show good agreement, with results being within 15%, for photon energies greater than 0.1 MeV, i.e. photon energies of importance to radionuclides found in the environment.

Figure 4.2-1 shows a layer of contaminated material of finite thickness (L_2-L_1) and infinite lateral extent, buried beneath a thickness of material which may or may not be of the same material type. To calculate photon dose rate the GRANIS model calculates total photon flux density from such a layer by first considering the flux density from an annulus of a simple disc at the top of the contaminated layer and then integrating over all disks within the layer.

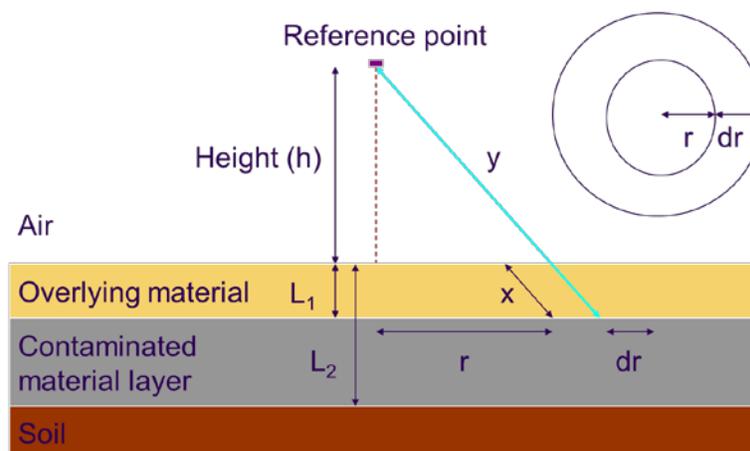


Figure 4.2-1: Schematic of GRANIS model application

The relationship between absorbed dose rate in air, and the particle flux density for mono energetic radiation of energy under conditions of electron equilibrium (Goussev, 1968) is determined by:

$$D = P \cdot E \left(\frac{\mu_a}{\rho} \right)_{air} \cdot \Psi \quad (4.21)$$

Where:

- D = absorbed dose rate in air (Gy h⁻¹)
P = constant (9.6 10⁻¹² J MeV⁻¹ s h⁻¹).
ψ = particle flux density (photons s⁻¹ m⁻²)
(μ_a)_{air} = mass attenuation coefficient of air (m² kg⁻¹) for photons of energy E
ρ = density of air (kg m⁻³)

Note: the above equation is accurate up to a gamma radiation energy of 3 MeV, at higher energies electron equilibrium is difficult to attain giving approximate dose estimates.

From (Goussev *et al.*, 1968), for a disk source with a slab shield, i.e. the disk of activity at the top of the contaminated layer as shown in Figure 4.2-1, shielded by the slab of contaminated material above it, the radiation flux at a point on the centre line of the disk source may be obtained by the integration of the attenuation function for point sources which make up an annulus of the disk. The flux from the differential ring source (annulus) contained between the circles of radii r and r + dr is:

$$d\Psi = B \cdot C_A \left(\frac{2\pi r dr}{4\pi y^2} \right) e^{-\mu_s x} e^{-\mu_a(y-x)} \quad (4.22)$$

Where:

- B is the build-up factor for scattered radiation.
C_A is the number of photons of initial energy E (MeV), emitted per second per m² of surface and per cm depth of soil in the layer considered.
x, y and r are lengths (metres) indicated in Figure 4.2-1.
μ_a and μ_s are the linear attenuation coefficients for the photons of energy E in air and material respectively (m⁻¹).

Note on build-up: The effects of photon scattering which occur in the material of the contaminated layer, shielding layer and the air layer, which scatter photons from their initial trajectory away from the reference point, back to the reference point are accounted for though the use of build-up factors. The amount of scattering and thus the size of the build-up factor for each annulus is dependent on the shield material, and naturally varies with photon energy and attenuation depth of the annulus. Generally, build-up is higher for low atomic number materials. Combining the two equations above, the following is obtained:

$$dD(r, l, E) = \frac{P \cdot E}{2} \left(\frac{\mu_a}{\rho} \right)_{air} B \cdot C_A \left(\frac{r dr}{y^2} \right) e^{-\mu_s x} e^{-\mu_a(y-x)} \quad (4.23)$$

Where:

- dD(r,l,E) is the component of dose in Gy per year due to photons of energy E MeV from an annular source between r and r+dr, at a depth of l in the material (see Figure 4.2-1).
l is the depth of a disc annulus.

To calculate the exposure from a whole layer of contaminated material the dose from a potentially buried annulus requires integration over the radius of the annulus ($r = 0$ to infinity) and over the vertical thickness of the layer (L1 to L2 in Figure 4.2-1). This integration is performed in the GRANIS code.

4.2.2 ECOFOR

ECOFOR is a model designed to investigate the transfer of radionuclides from an initial atmospheric deposition or from an underground source to vegetation (the model is set-up for pine trees) followed by further translocation to the various structural components of the tree and cycling back to the ground. ECOFOR aims to simplify the hydrological problem by assuming a quasi-steady-state of laminar water flow in a 10-layer soil column of heterogeneous composition where each of the 10 soil layers S_{ij} is allowed to input or output to/from adjacent layers, depending on the degree of water saturation.

In this project, ECOFOR was improved and adapted to the Belgian NORM site. A basic schematic of the model is given in Figure 4.2-2. A simplified approach for the hydrology is used, which superimposes a 'tipping bucket' approach (in which soil layers fill as adjacent layers become filled with water), with simplified dynamic equations (Darcy law and Lucas-Washburn capillarity equations) used to calculate the flow of water and solutes. This is one of the main differences between ECOFOR and TALAL (see Section 4.2.3), a SVAT model that has been adapted for the Belgian NORM site during this project, which simulates soil moisture and solute dynamics using an integrated approach that couples Richards equation to an advection-dispersion equation.

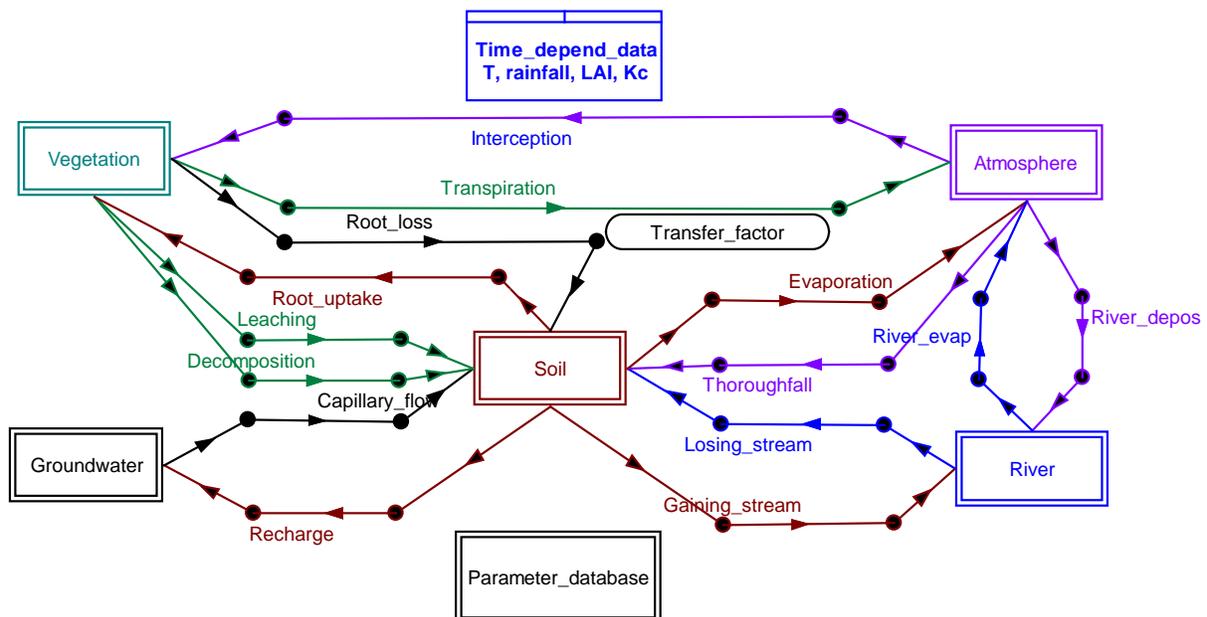


Figure 4.2-2: A basic description of the ECOFOR model

4.2.2.1 Soil water and element transport

ECOFOR calculates water transport between soil layers mediated by a system of “tipping bucket switches”. The volumetric water content in each layer can fluctuate between a minimum, defined by the soil field capacity, and a maximum, defined by the soil porosity, until the resulting water table fills the whole soil pore volume, leading to runoff (Ohashi, 2017).

Regarding the hydrological processes, the vertical flow from above into deeper soil regions is simplified as taking place mainly via advection, which is expressed with the Darcy down-flow equation:

$$\phi_{ij}^{down} = \frac{T_j R_{ij}}{z_i \varepsilon_i} S_{ij} \text{ if } i = 1 \quad (4.24)$$

$$\phi_{ij}^{down} = \left(1 - k_{i+1}^{sat}(\theta_{i1})\right) \frac{Hc_i \cdot SA}{z_i} \text{ if } i = 2 \dots 10 \quad (4.25)$$

With i being the index of the vertical layers (number one is the uppermost), T_j as throughfall (mm d^{-1}), R_{ij} the retardation factor, z_i the thickness of the soil layer (m), ε_i the porosity, S_{ij} the content of soil layer i (m^3), θ_{i1} the volumetric water content, Hc_i the hydraulic conductivity layer (m s^{-1}), SA the surface area of soil column (m^2) and $k_{i+1}^{sat}(\theta_{i1})$ a purposely-defined saturation switch which is 1 by default and 0 if the layer is saturated.

The upper capillary transport is described by Newton dynamics' equation for a viscous non-compressible liquid, assuming quasi-steady Poiseuille flow.

4.2.2.2 Uptake by the tree

The plant root water uptake rate r_{ij}^{uptake} is a time-dependent variable, and in optimum conditions it is linked to the potential (maximum) evapotranspiration demand which is modelled with the Monteith equation (Monteith and Unsworth 2007):

$$ET_0 = \frac{\Delta(R_n - G) + \frac{\rho_a c_p (\delta e)}{r_{av}}}{\lambda_v \rho_w \left[\Delta + \gamma \left(1 + \frac{r_s}{r_{av}} \right) \right]} \quad (4.26)$$

Where Δ is the rate of change of saturation specific humidity with air temperature (slope vapour pressure curve) [Pa K^{-1}]; $R_n - G$ is the net energy input from solar radiation minus soil heat storage, ρ_a = dry air density [kg m^{-3}], c_p is the Specific heat capacity of air [$\text{J kg}^{-1} \text{K}^{-1}$], $\delta e = e_s^0 - e_a = (1 - \text{relative humidity}) \times e_s^0$ with e_s^0 being the mean saturated vapour pressure [Pa]; r_{av} = is the bulk surface aerodynamic resistance for water vapour [s m^{-1}]; γ = is the psychrometric constant; λ_v = is the latent heat of vaporisation; r_s the reference crop surface resistance [s m^{-1}], and ρ_w is the density of water [kg m^{-3}].

The root uptake itself occurs through a system of branching roots that can be represented by an exponential root water uptake model. It considers root uptake balanced against evapotranspiration, xylem upflow, and phloem downflow along the osmotic pressure gradient. The used exponential root water uptake model is based on the assumption that the root has an exponential distribution with depth. Furthermore, soil drying (wilting point) and waterlogging (anaerobiosis, using a water stress coefficient) are considered. The vertical flow within the tree is modelled as the ascent of xylem across a hydraulic potential gradient, and the phloem transport from leaves to root is described using the Poiseuille equation

with a downward-directed osmotic pressure. The movement of solutes with respect to water in the soil compartments is assumed proportional to retardation, which relates to the empirical parameter K_d . In the plant, selectivity coefficients are applied to calculate the retardation of solutes with respect to the water fluxes, as in a previous modelling approach. From here, it is assumed that radionuclide fluxes can be coupled to element analogue fluxes (Casadesus, *et al.* 2008).

The main ECOFOR parameters are described elsewhere (Vives I Batlle *et al.* 2014). The key and most sensitive parameters are the soil K_d , porosity, field capacity and saturated hydraulic conductivity. ECOFOR has undergone a great deal of improvement since its inception, particularly in the present project. The main improvements compared with previous versions are (a) the discretisation of the soil column into a multi-layered, heterogeneous soil profile, (b) improved representation of the unavailable and available fraction of radionuclides in soil, (c) a more compact representation of the hydrology balancing hygroscopic, capillary and gravitational water, and (d) improvement of the tree sub-model with sap transport and linkage of elements by means of selectivity coefficients. Whilst being considered “complex” and process-based, ECOFOR not as sophisticated as some forest models that include plant biology and chemical processes at the molecular level (Deckmyn *et al.*, 2011). Therefore, it represents a research model of the type “sufficiently complex to be realistic and sufficiently simple to be practical” capable of future use in environmental assessments.

4.2.3 TALAL

This model has been developed to simulate water and solutes – including naturally occurring radionuclides (NORM) - dynamics in the soil-plant-atmosphere (SPA), and it is based on the well-established knowledge of water flow and solute transport in soils as well as uptake by plant roots (this modelling approach is widely used in reactive transport modelling).

The TALAL model combines a process-based soil module and an empirical tree module to simulate radionuclides behaviour in a forest ecosystem under a wide range of environmental conditions (e.g. weather, soil profiles, tree species). While the soil module simulates water and solute dynamics in the root zone, the tree module - of the type commonly used in radioecological studies (Van den Hoof and Thiry 2012) - simulates cycling of substances within a tree. Like ECOFOR, this model was originally verified, calibrated and validated using data from the experimental Scots pine forest in Mol and has now been applied to the Belgian observatory site.

The model accounts for the following processes:

- Infiltration of atmospheric water into soil, subsequent percolation and deep drainage (e.g. recharge into aquifer),
- Capillary rise from the local water table and upward flow in the soil column,
- Redistribution of moisture over time and space,
- Water uptake by plant roots subject to water availability (accounts for stress due to drought),
- Mass flow of soil solutes,
- Molecular diffusion of solutes due to concentration gradients in soil,
- Hydrodynamic dispersion of solutes due to heterogeneity of advection velocities,
- Concentration and dilution of solutes due to changes in moisture,

- Retardation due to sorption,
- Interception of atmospheric deposition by plant foliage,
- Active root uptake of soil solutes in proportion to plant root profile and the absorption power of the roots,
- Solute cycling in plants (currently three compartments: roots, trunk and foliage),
- Feedback loop of solutes from tree to litter (through litterfall),
- Feedback loop of solutes from litter and decomposing roots to soil,
- First order decay processes (e.g. radioactive decay)

In TERRITORIES, TALAL model is applied to simulate short- and long-term dynamics of NORM in the root zone and subsequent transfer to trees. The model calculates a wide range of outputs (e.g. time series and profiles of soil moisture and solutes, leaching fluxes to deep soil and groundwater, concentrations in vegetation and proportional distribution in different tree parts). In the current application, the model is adapted to calculate soil-to-tree transfer factors for different NORM for the sake of comparison with simple, empirical models and with measurements when available.

The following paragraphs describe the mathematical specification of the TALAL model. Some processes in the model were removed because they were irrelevant to the study site (e.g. interception of NORM by tree foliage).

4.2.3.1 Soil module

In TALAL, the soil module simulates water flow and solute transport and subsequent uptake into plants via their roots. The vertical soil water flow is simulated according to the one-dimensional physical model of water flow in porous media (Richards equation) (Richards, 1931):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left(K \frac{\partial(\varphi+z)}{\partial z} \right) + \phi \quad (4.27)$$

Where:

- θ ($\text{m}^3 \text{m}^{-3}$): the volumetric water content,
- z (m): soil depth (directed positive downward),
- t (day): time,
- K (m s^{-1}): soil hydraulic conductivity,
- ϕ (s^{-1}): the root water uptake rate.

K and θ are highly nonlinear functions of φ that are described using van Genuchten function (Van Genuchten, 1980) functions.

The model uses a φ -based form with an iterative solver (continuity in layered profiles and under saturated conditions). A backward Euler integration scheme is adopted as it gives the highest stability, the model uses adaptive time stepping for greater speed and a variable spatial grid in order to improve computational efficiency.

The module simulates the vertical transport and distribution of NORM using a standard advection-dispersion equation with a sink/source term:

$$\frac{\partial(\theta Rc)}{\partial t} = -\frac{\partial(qc)}{\partial z} + \frac{\partial}{\partial z} \left(D \frac{\partial c}{\partial z} \right) + S \quad (4.28)$$

Where:

- c (Bq m⁻³): solute concentration in porewater,
- q (m d⁻¹): darcy flux,
- D (m² d⁻¹): dispersion coefficient
- S (Bq m⁻³ d⁻¹): solute sinks and/or sources in soil.

The soil module is flexible and customisable. It could readily be adapted to simulate soil profiles with contrasting hydraulic or sorption characteristics. Such contrast may result from changes in soil texture, porosity or organic matter content. For instance, coarse-textured soils (e.g. sand) tend to have greater hydraulic conductivity than fine-textured soils (e.g. clay). Organic soil horizons tend to accumulate certain radionuclides compared to mineral horizons. The model could simulate rather complex, multilayer soil profiles through proper discretisation (i.e. number of the numerical layers the model uses) and parameterisation of the model.

4.2.3.2 Tree module

Once taken up into the tree, the tree module cycles the NORM within and between the tree compartments. The tree module comprise a system of first order differential equations each of which describe the activity balance of the NORM within a given tree compartment. Currently, the module includes three compartments: the roots, wood (trunk) and foliage. The inter-compartmental transfers of the NORM are expressed in terms of constant rate coefficients.

In conclusion, the principal way in which TALAL is more “advanced” than ECOFOR is in adopting the dynamically-coupled Richards equation and advection-dispersion equation system for water and the radioelements. In this project, we took the tree transfer factor as its endpoint for comparing model outputs and we did not consider plant interception or translocation due to the low values of measurements in plant parts at the Belgian NORM site (generally close to detection limit, meaning that most radionuclides are concentrated in the roots). Therefore, for the purposes of this study, when comparing the two SCK-CEN models, we are comparing the more advanced water and solute transport features of the TALAL model with the simplified representation of the hydrology considered in ECOFOR.

4.2.3.3 Model set-up

Data used

TALAL parameterisation for the NORM site

The model requires various inputs and parameters, which could either be measured directly or estimated from other readily available measurements. The model inputs could broadly be grouped into weather-specific, soil-specific and plant-specific.

Weather-specific inputs include measurements of air temperature, rainfall, wind speed, relative humidity and solar radiation, which had been reported for the study region in earlier work (Ohashi, 2017). Potential evapotranspiration (ET_p), the sum of evaporation and plant transpiration, is estimated according to the method of Penman-Monteith (Monteith and Unsworth 2007) and subsequently partitioned into potential evaporation and transpiration using Beer’s law.

Soil-specific inputs included the key physical and chemical characteristics of the clay-like CaF_2 sludge. The model was set up using hydraulic data appropriate for clay (Table 4.2-1). TALAL was setup to simulate the upper 60 cm of the sludge, which is the depth for which we have site-specific measurements and where most tree roots are typically found. Table 4.2-1 and Table 4.2-2 present the physical and chemical parameters used in the model.

Table 4.2-1: Physical characteristics of the CaF_2 sludge profile as defined in the TALAL model. The hydraulic characteristics are typical of clay soil.

Soil layer	Bulk density [g/cm ³]	Residual water content [m ³ m ⁻³]	Saturated water content [m ³ /m ⁻³]	Saturated Hydraulic conductivity [m d ⁻¹]
0 – 20 cm	2.1	0.1	0.50	0.26
40 – 60 cm	2.4	0.1	0.50	0.25

Table 4.2-2: Sludge-specific solid-liquid distribution coefficient (Kd) for NORM.

Soil layer	Kd [cm ³ /g]
Barium	463
Lead	128
Uranium	306

Plant-specific inputs included leaf area index, root depth profile, a crop factor and a set of soil matrix potential values to model tree transpiration and water uptake processes and their response to water stress. Because measuring these parameters was not foreseen in this project, we used (same as ECOFOR) values in published studies (Ohashi, 2017; Brission, 1992). To calculate the NORM activity concentration in tree compartments, the biomass of the root, wood and foliage compartments were needed. The biomass of the trunk has been reported (Gielen, Vives i Batlle *et al.* 2016) for an average pine tree in the region. We calculated the root biomass using allometric relations for coarse and small roots of Scots pines (Xiao and Ceulemans 2004).

4.2.4 TREE4-advanced

Significant improvements have been made to the TREE4 model during the last 2 years. First of all, efforts have been undertaken to refine the description of the structure and functioning of the soil-tree subsystem through modifications made in the conceptual model. Much work has also been done to improve the mathematical parameterisation of some transfer processes, whose kinetic rates (in d⁻¹) now depend on some driving eco-physiological, hydrological or soil characteristics. As indicated in Figure 4.2-3, two newly developed modules were implemented in TREE4-advanced, in order to model: (i) the hydrological cycle within the system and (ii) the eco-physiological characteristics of the forest stand and their annual evolution (i.e. ageing). These improvements are briefly described hereafter. The calculations performed by the improved Cs transfer module now rely on the estimation of the annual growth/decay of some eco-physiological characteristics such as the forest stand density, the stem volume or the mean tree height (see hereafter). The tree species and the age of the forest stand at start date must be specified. The calculation also relies on the prediction of the mean annual water fluxes such as evapotranspiration fluxes or water uptake by vegetation, based on the mean annual climatic characteristics.

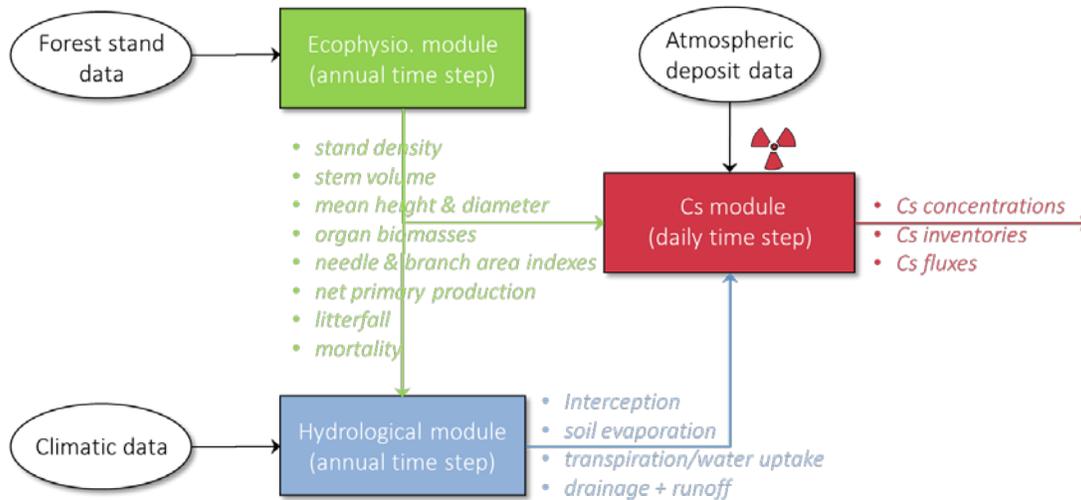


Figure 4.2-3: TREE4-advanced model structure. See main text for description.

Cs module

The refinement of the conceptual model, as depicted in Figure 4.2-4, mainly consisted of disaggregating some tree and soil compartments, discretizing the mineral soil profile into thin horizontal layers while explicitly accounting for convection/dispersion mechanisms, and introducing new transfer processes, the importance of which was revealed in the months or years following the Fukushima accident. In this advanced approach, we now differentiate the canopy into two distinct sub-compartments, i.e. the foliage and the branches. Each of them is in turn splitted into an “external” pool, directly impacted by the atmospheric fallouts, and an “internal” pool contaminated by foliar incorporation, root uptake or internal translocation processes. Unlike for the bark compartment, no distinction is made between branch wood and trunk wood because we assume that ^{137}Cs concentration in wood is homogeneous. Foliar excretion and in-bark immobilization processes were introduced in the model in order to sustain ^{137}Cs contamination in throughfall waters on the long term, as we observed it in Fukushima forests. The long-term decrease of the total ^{137}Cs inventory in (living) trees due to natural mortality or tree harvesting can now be accounted for because time evolution of the forest stand density is explicitly predicted. Within the “organic” layer, we now consider a so-called “non-leachable” pool. This pool is continuously enriched by litterfall but progressively decomposed into a “leachable” pool as a result of organic matter decomposition. In the mineral soil profile whose characteristics can vary with depth, we consider that the vertical transport of ^{137}Cs is governed by classical convection-dispersion mechanisms. This layer is discretised into a series of thin sub-layers of typically 0.5 cm depth, inside which bio-available/leachable ^{137}Cs can be reversibly fixed onto/remobilised from the solid matrix.

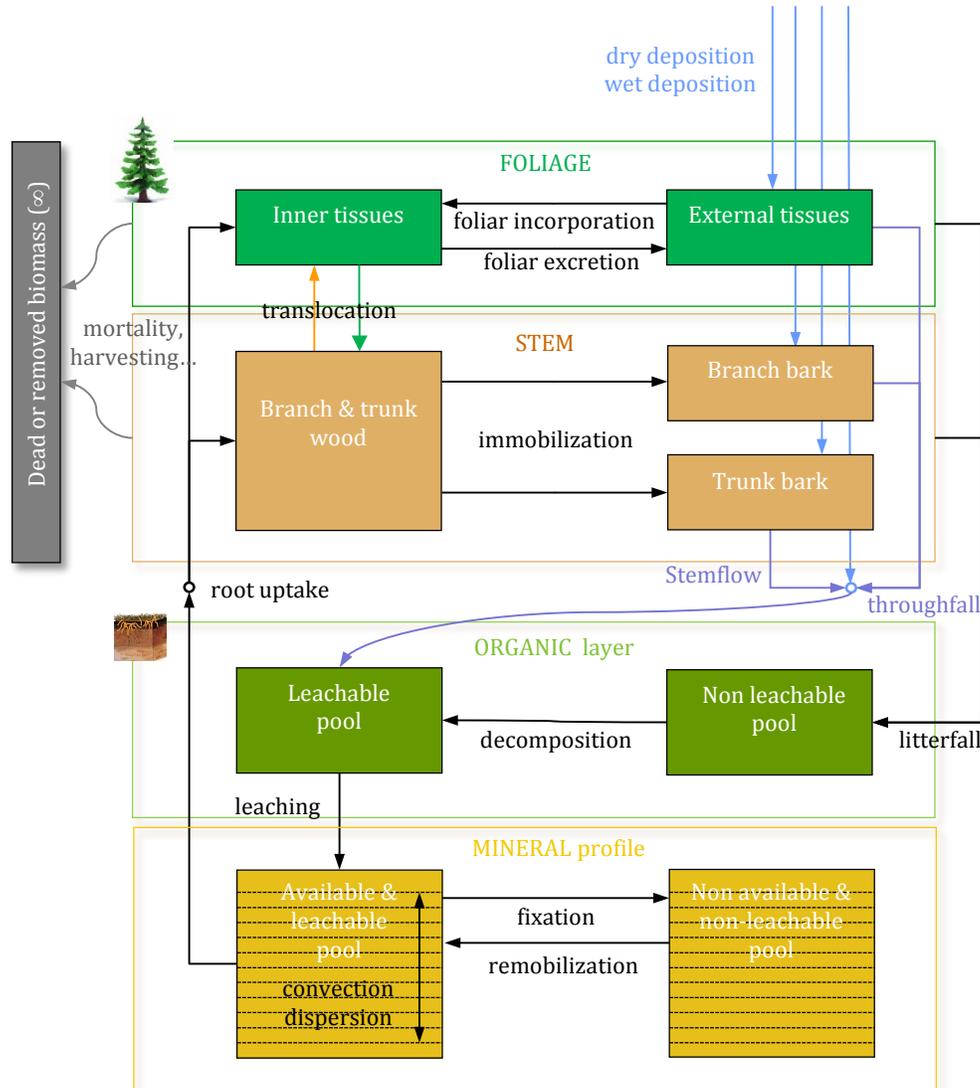


Figure 4.2-4: Conceptual modelling of Cs cycling within the soil-tree system (TREE4-advanced). See main text for description

Like in the TREE4-simple version, the mathematical approach adopted for predicting ^{137}Cs inventories and fluxes relies on a system of non-stationary differential equations describing activity conservation in the soil-tree system. Concentrations are deduced from inventories using (bio) masses. While bulk densities are imposed for soil layers, tree organ biomasses and their annual evolutions are now explicitly predicted by the eco-physiological module. Regarding the parameterisation of transfer processes, many improvements have been introduced, as listed in see Table 4.2-3. To summarise, most of the transfer processes related to tree vegetation now depend on biomass and K fluxes, because we assume an analogy between K and Cs like in the BioRUR approach (Casadesus *et al.*, 2008). The migration processes driven by the movement of water, such as throughfall, stemflow and convection, now depend on the annual water fluxes estimated by the hydrological module.

Table 4.2-3: List of the transfer processes for which the mathematical parameterisation has been improved or newly introduced in TREE4-advanced (the latter are marked by an asterisk), with indication of the eco-physiological, hydrological or soil properties the transfer rates now depend on.

Transfer process	TREE4-simple parameterisation	TREE4-advanced parameterisation
Root uptake	Constant rate	Rate dependent on : annual K demand for tree biomass growth, annual transpiration rate
(*) Foliar excretion	-	Rate dependent on : annual K excretion by foliage, K content in living and dead foliage
Litterfall	Constant rate	Rate dependent on : annual turnover rate of foliage and wood biomasses, K content in living and dead foliage
Throughfall	Constant rate	Rate dependent on : annual rainfall, foliage area index, branch area index
Stemflow	Constant rate	Rate dependent on : annual rainfall, trunk area index
(*) Immobilisation	-	Rate dependent on : annual K demand for branch and bark growth, K content in bark
(*) Mortality, thinning	-	Rate dependent on : annual mortality rate
(*) Translocation	-	Constant rate
(*) Decomposition	-	Rate dependent on : OM decomposition rate
Leaching	Constant rate	Rate dependent on : annual percolation rate
(*) Convection/dispersion	-	Rate dependent on : annual transpiration flux, annual drainage flux, soil water content, dispersivity coefficient
(*) Remobilisation	-	Constant rate

Such a complexification of the conceptual and mathematical approaches significantly increases the number of model parameters required, although a majority of these physical, biological and chemical parameters can be easily estimated from literature or directly measured at field.

Eco-physiological module

This module relies mainly on the use of empirical relationships which were established from field observations throughout Middle Japan for the two types of even-aged coniferous plantations considered in our study. First of all, the forest stand is characterised by the following age-dependent variables: stand density (number of trees per m^2), mean trunk diameter at breast height (m), dominant tree height (m) and trunk volume (m^3 trunk wood per m^2). In case of a managed plantation, it was demonstrated that their evolution with age obeyed a Mitscherlich law, with coefficients values specific to the quantity considered. At any age, the biomasses of the four tree organs considered – foliage, trunk wood, trunk bark and branches (kg dry mass m^{-2}) – are deduced from the predicted trunk volume (or trunk diameter) using allometric relationships. As indicated in Table 4.2-3, the parameterisation of the transfer processes related to tree vegetation further requires knowledge of the biomass fluxes (kg dry mass $m^{-2} y^{-1}$), i.e. net primary production, mortality rate and biomass turnover rate for each of the four organs. These are estimated by solving each year a mass balance equation. Finally, area indexes ($m^2 m^{-2}$) for foliage, branches and trunk

are estimated from the corresponding predicted biomasses, using specific area coefficients ($\text{m}^2 \text{kg}^{-1}$ dry mass) which must be specified. These are very uncertain for woody organs because their (effective) values strongly depend on the characteristics of each rainfall event and must account for shielding effects within the vegetation layer.

Hydrological module

The hydrological module calculates the following annual water fluxes (mm y^{-1}): interception fraction by the above-ground biomass, soil evaporation, tree transpiration and drainage to the groundwater (run-off was neglected in our study). The first three contributions are estimated through the use of empirical relationships which were established from field observations for a variety of forested watersheds throughout Middle Japan. These relationships involve 2 driving variables: the annual temperature and the annual precipitation recorded at each site. Drainage is estimated from the three other contributions by solving a simple mass balance equation in which water storage is neglected. This module further estimates the vertical profile of the water percolation flux within the mineral horizon, taking into account an idealized root density profile (to be specified).

4.2.5 AMIS

To better describe the marine environment near the Sellafield facilities the local model for the Irish Sea regions (AMIS) has been developed. The model is a modification of the NRPA box model. The AMIS model uses the same modified approach for the box modelling described in the Section 4.1.5. Surface structure of the local model is presented in Figure 4.2-5. The structure of the boxes in the model AMIS for the Irish Sea is similar to the NRPA box model (ARCTIMAR). In addition, the model includes the surrounded box “Rest of seas”. The intertidal beach area is part of the Cumbrian Water box (the red star) is also shown in Figure 4.2-5.



Figure 4.2-5: The structure of the surface water boxes for the AMIS box model (30 compartments).

4.2.5.1 The choice of the environmental parameters

Based on the published information about the parameters used in the AMIS model, their values can vary up to 2-3 orders of magnitude (MacKenzie and Nicholson, 1987; Nielsen *et al.*, 1997; Mitchell *et al.*, 1999; Iosjpe *et al.*, 2002; MARINA II, 2003; IAEA, 2004; Perianez, 2003 & 2005).

The liquid discharges of the radionuclides Cs-137, Pu-239 and Am-241 have been selected for the potential improvement of the AMIS model parameters by comparing the results of the simulations with the available monitoring data for the “Cumbrian Waters” box.

It has been previously shown that the choice of appropriate model parameters has a powerful impact to the results of the radioecological analysis (Iosjpe, 2011A). The selection of parameters has been performed with the help of a sensitivity analysis (similar to Iosjpe (2011, 2011A) with the following sensitivity index (Jørgensen, 1994):

$$S^{(L)}(P) = \frac{dP^{(S)}}{dP} \frac{P_0}{P_0^{(S)}}, \quad (4.29)$$

where $P^{(S)}$ and P correspond to state parameters (for example, concentrations of radionuclides in water and sediment phases, doses to man and biota, etc.) and parameters which are under evaluation; P_0 and $P_0^{(S)}$ correspond to the basic values.

According to expression (4.29): $S^{(L)} > 0$ when the state parameter $P^{(S)}$ increases with the increase of the evaluated parameter P . $S^{(L)} < 0$ when $P^{(S)}$ decreases with the increase of the parameter P . There is no influence of the parameter P to the state parameter $P^{(S)}$ when $S^{(L)} = 0$.

Figures 4.2-6 and 4.2-7 show examples of calculations of the local sensitivity indexes for the Cs-137 concentrations in the filtered water for parameters reworking rate (Rw), sedimentation rate (SR) and apparent sediment distribution coefficient (k_d). All calculations correspond to liquid discharges of radionuclides into the Cumbrian Waters.

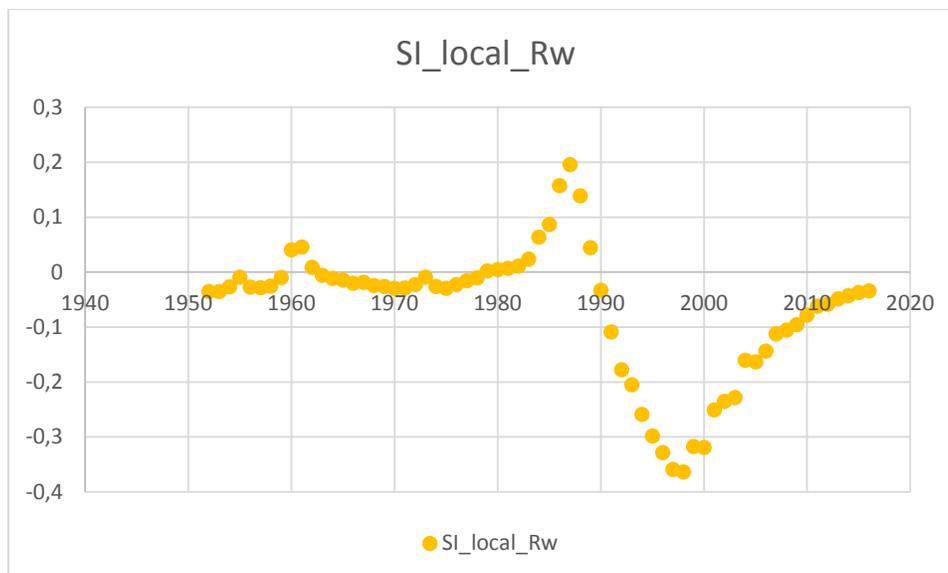


Figure 4.2-6: Dynamic of the local sensitivity index for the reworking rate.

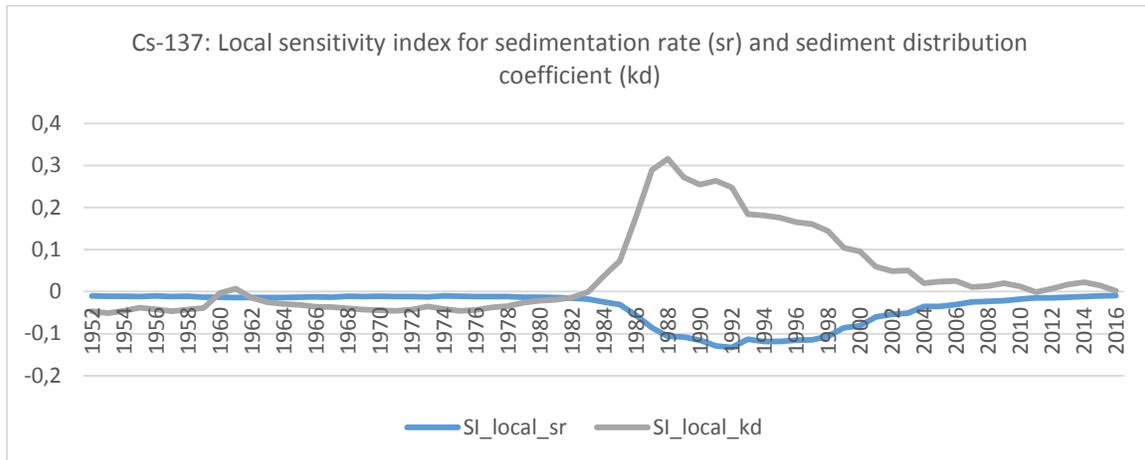


Figure 4.2-7: Dynamic of the local sensitivity indexes for the sedimentation rate and sediment distribution coefficient.

The simulations presented in Figures 4.2-6 and 4.2-7 clearly demonstrate the complexities encountered when modelling dispersion of the radioactivity in the marine environment. The results show that the radionuclide concentration can either increase or decrease with the increase of the evaluated parameters. It is also shown that the results can strongly depend on the time of analysis.

Use of the sensitivity indexes improves our knowledge about the influence of parameters to the model end points.

The most suitable set of model parameters have been selected for all three radionuclides (Cs-137, Pu-239 and Am-241) simultaneously according to the following expression:

$$\min_{\{\vec{P}_\mu\}} \left\{ \sum_{l=1}^3 \sum_{i=1}^{n^{(l)}} \left[C_i^{(\vec{a}_\mu, k_d^{(\mu, l)})} - \hat{C}_i^{(l)} \right]^2 \right\} \quad (4.30)$$

Here $\{\vec{P}_\mu\} = \{\vec{a}_\mu, k_d^{(\mu, l)}\}$ are the sets of different model parameters, which have been used for parameterization ($\mu=1, \dots, M$, where M is the number of different sets in the present investigation); each set of parameters \vec{P}_μ consists of the set of environmental parameters \vec{a}_μ and the set of “apparent” sediment distribution coefficients $k_d^{(\mu, l)}$ ($i= 1, 2, 3$ for radionuclides Cs-137, Pu-239 and Am-241, respectively); $i=1, \dots, n^{(l)}$, where $n^{(l)}$ is number of measurements of the concentration of radionuclide l in water and sediment; $C_i^{(\vec{a}_\mu, k_d^{(\mu, l)})}$ and $\hat{C}_i^{(l)}$ are concentrations of radionuclide i in the water and sediment phases, calculated by the model, and experimental data, respectively.

It is necessary to note some important points concerning the sediment distribution coefficients. The definition of the sediment distribution coefficients (k_d) is based on assumptions about the equilibrium balance between dissolved and particulate phases (IAEA, 2004). This assumption is not generally supported by the real conditions in marine environments (Periáñez *et al.*, 2018). Therefore, terms “site-specific” and “apparent” k_d are used in some investigations (Iosjpe, 2011A, Periáñez *et al.*, 2018). Kinetic sub-models for the exchange of radionuclides between water and sediment phases require k_d to be under equilibrium

conditions in order to define the system of kinetic coefficients (Periáñez, 2003, 2005; Periáñez *et al.*, 2018). Additionally the kinetic sub-models can construct “apparent” k_d during numerical simulations. In particular, it was shown, that (i) “apparent” k_d value near the source of contamination can be 2-3 times less than the equilibrium value and (ii) apparent k_d value in the sediment can be 10-1000 times less than the equilibrium value (Periáñez *et al.*, 2018).

5 CASE STUDY: NORWEGIAN FEN FOREST SITE

5.1 Introduction to the case study

The Fen Complex is located in the Telemark County in south-eastern Norway. The area has for many decades been subject to public and mining exploration activities due to the presence of large estimated quantities of thorium (Th) ore that according to Berg *et al.* (2012) could reach 675 000 tonnes. However, although rocks rauhaugite and rødbergite are known to be particularly rich in ^{232}Th , the uncertainty about the exact quantities has been emphasized and further work on estimations is needed (Thorium Committee, 2008). Apart from ^{232}Th , the Fen Complex has been an area of interest, for many years, due to mining of iron (Fe) and niobium (Nb), as well as due to possibilities for extraction of rare earth elements (REE). High levels of naturally occurring radioactive material (NORM) in rocks and soil with heterogeneous distribution and observed ‘hot spots’ were measured previously in the Fen Complex (IFE, 2006, NGI-NMBU; 2010, Mrdakovic Popic *et al.* 2011). The area of interest is shown in Figure 5.1-1.



Figure 5.1-1: Map of Fen complex showing locations of the Sjøve and Fen mining areas and Lake Nordsjø (blue area). Reproduced from NGI-UMB (2010).

The Fen Complex data, for inclusion in the TLD and subsequent analysis in this report, were collected during fieldwork organized in period 2008-2011 as a part of PhD and MSc projects at Norwegian University of Life Sciences (NMBU) and as a part of activities of Centre for Environmental Radioactivity in Norway (CERAD). Data that have been provided for the TLD, primarily from Mrdakovic Popic *et al.* (2011; 2012)

and Mrdakovic Popic (2014), include terrestrial gamma dose measurements, soil, plants and earthworms activity concentrations of ^{232}Th , ^{238}U and certain decay-chain progeny and results of basic soil chemistry analyses.

Terrestrial gamma dose rates in the outdoor air were measured in several fieldwork campaigns, during all seasons to account for possible seasonal variation. All measurements were made at 1m height above the ground. A portable gamma detector (Geiger Müller counter) type Automess (radiacmeter 6150 AD 4 LF) with a response range 0.01 – 20.00 $\mu\text{Gy}/\text{was}$ was used. Hand measurements were organized at sub-sites delineated by regular grids 10 x 10m while readings were repeated until a constant signal was obtained. In addition, thermo-luminescent dosimeters (TLDs) type MR 200C with luminescence material CaF_2 doped with Mn, developed by the Jozef Stefan Institute, Slovenia, were used for a long term measurement (3 months) of gamma radiation.

Samples of soil, plant and earthworms were collected during different fieldworks in period 2008-2011, in different seasons. Some samples for analysis of the vertical distribution were collected at depth of 0.05, 0.1, 0.2, 0.25 and 0.3 m by incrementally excavation of the wall of the soil pit, but this was limited due to high abundance of the stones in the pits of soil. Basic soil physical and chemical parameters (pH, grain size, H_2O (%), OM (%), P_2O_5 , CaCO_3 , CEC, Exch, Ca) were determined. Some soil samples were analysed gamma spectrometry whilst others were determined by ICP-MS.

Plants and earthworms were collected at different sampling points within sites. Plants that found to be in abundant were regarded as representative, and in total, nine different plant species (moss, lichen, fern, birch, pine, spruce, dandelion, grass, strawberry) and four earthworms (endogeic and epigeic) species were collected. The chosen plants covered a relatively wide range of wild forest flora, with vascular and non-vascular plants, flowering plants, deciduous and coniferous trees. Earthworms were collected by digging the soil and hand sorting of organisms that were then transported to laboratory in the plastic boxes together with soil, roots and leaf litter. After identifying the plant species, aboveground plant parts and root were divided, cleaned and dried. Acid microwave decomposition was done according to standard protocols and radionuclides were measured by ICP-MS. In the same way as plants, earthworm species were identified after acid-microwave decomposition were measured by ICP-MS.

The data set described above has been supplemented in the TLD by an extensive gamma air kerma rate survey undertaken, in 2010, over a high-resolution sampling grid for the Sjøve mining complex in Fen and including a limited number of additional soil measurements. These data were collated as part of the Norway Grants “PORANO” project and have been reported in Dowdall *et al.* (2012).

5.2 Application of the models to case study

5.2.1 CROM

We used the data from Fen site as shown in Table 5.2-1 to compare dose rates obtained using simple dose conversion factors (DCFs) implemented in the CROM code, calculated with the IAEA SRS 19 (IAEA, 2001) assumptions and the unit conversion factors derived from the FGR12 report (EPA, 1993), which convert radioactive deposit in a soil (Bq m^{-2}) to equivalent ambient dose (Sv y^{-1}), with directly measured equivalent ambient dose *in situ*.

Table 5.2-1: Activity concentrations of the natural radionuclides measured in the soils in Fen ($Bq\ kg^{-1}\ fw$)

U238	Ra226	Pb210	Th232	Ra228	Th228	K40
1.17E+03	1.17E+03	1.17E+03	1.55E+03	1.55E+03	1.55E+03	4.06E+02
3.09E+01	3.09E+01	3.09E+01	5.53E+01	5.53E+01	5.53E+01	4.06E+02
9.70E+02	6.60E+01	6.60E+01	1.01E+04	1.01E+04	1.01E+04	4.06E+02
4.10E+02	9.60E+01	9.60E+01	3.38E+03	3.38E+03	3.38E+03	4.06E+02
9.02E+01	9.02E+01	9.02E+01	1.77E+03	1.77E+03	1.77E+03	4.06E+02

The units of these DCFs must be converted because the measured activity concentrations are measured in $Bq\ kg^{-1}$, the ambient dose equivalent is measured in $\mu Sv\ h^{-1}$ and the DCFs relate ambient dose equivalent in $Sv\ y^{-1}$ with surface deposit in $Bq\ m^{-2}$. The density of the soil used for the conversion was $1300\ kg\ m^{-3}$, and the depth of the soil considered was 5 cm. Performing the conversions, the obtained DCFs for those radionuclides would be as shown in Table 5.2-2.

Table 5.2-2: DCFs for each of the measured radionuclides in situ

Radionuclide	DCF	
	$Sv\ y^{-1}\ per\ Bq\ m^{-2}$	$\mu Sv\ h^{-1}\ per\ Bq\ kg^{-1}$
U238	9.50E-10	7.04E-06
Ra226	5.24E-08	3.89E-04
Pb210	1.12E-10	8.30E-07
Th232	4.99E-08	3.70E-04
Ra228	7.08E-08	5.25E-04
Th228	4.43E-08	3.28E-04

A value of $240\ \mu Sv\ y^{-1}$ ($0.027\ \mu Gy\ h^{-1}$), derived from elsewhere (UNSCEAR 2000), has been added to account for the cosmic component of radiation in the calculations. This value is for sea-level but the error introduced would be minimal considering the fact that the Fen field sampling area lies at an elevation of approximately 200 m above sea level.

As results are given in terms of ambient dose rate equivalents, a conversion factor was applied to provide output in the same units as the measurements data (provided as gamma air kerma rates in $\mu Gy\ h^{-1}$). In the energy range 0.3 to 3 MeV, the ratio of the ambient dose equivalent to the air kerma, or $H^*(10)/K_{air}$ (in units of $Sv\ Gy^{-1}$), lies in the range from 1.31 to 1.13 (ICRP 1996). For the sake of simplicity, a conversion (corresponding to 1.5 MeV) of 1.15 has been used in the subsequent calculations.

A multiplication of the soil activity concentration per the correspondent DCF (assuming secular equilibrium for all the radioisotopes not measured) provides the ambient dose equivalent for each radioisotope, the sum of which would provide the total value. The latter can be compared with the measured values together with the uncertainty for each predicted value as summarised in Table 5.2-3.

Table 5.2-3: Calculated – predicted - values of $H^*(10)$ ($\mu\text{Sv h}^{-1}$) and K_{air} ($\mu\text{Gy h}^{-1}$) in the five locations where soil activity concentrations were provided, together with the measured kerma rates at the same locations.

$H^*(10)$ ($\mu\text{Sv h}^{-1}$)	Predicted K_{air} ($\mu\text{Gy h}^{-1}$)	Measured K_{air} ($\mu\text{Gy h}^{-1}$)
2.411	2.096	1.050
0.127	0.110	0.750
12.436	10.814	1.460
4.222	3.672	1.410
2.245	1.952	1.670

Comparisons between modelled and empirical data were performed using the guidance provided in chapter 3.6.3. The root mean squared logarithmic error (RMSLE) was used to provide a measure of model performance.

$$\text{RMSLE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log y_i - \log \hat{y}_i)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(y_i / \hat{y}_i))^2} \quad (5.1)$$

Where :

\hat{y}_i = modelled value; y_i = measured value

A RMSLE of 0.59 was calculated for CROM based on a comparisons of 5 coupled predicted (model output) and measured data points.

5.2.2 NORMALYSA

NORMALYSA has been used to predict the activity concentrations in plants and ambient dose equivalent at the Fen site based on measured soil activities as input. Identifying groups of co-located measurements for soil and plants (and for soil and ambient dose equivalent) was an essential first step.

In order to identify groups of measured samples that correspond to the same geographical location, use was made of the functionalities available in the Google Earth (GE) software. All available measured data were exported from Excel spreadsheets of TLD to the GE. Data were visually inspected and corresponding groups identified as exemplified in Figure 5.2-1.

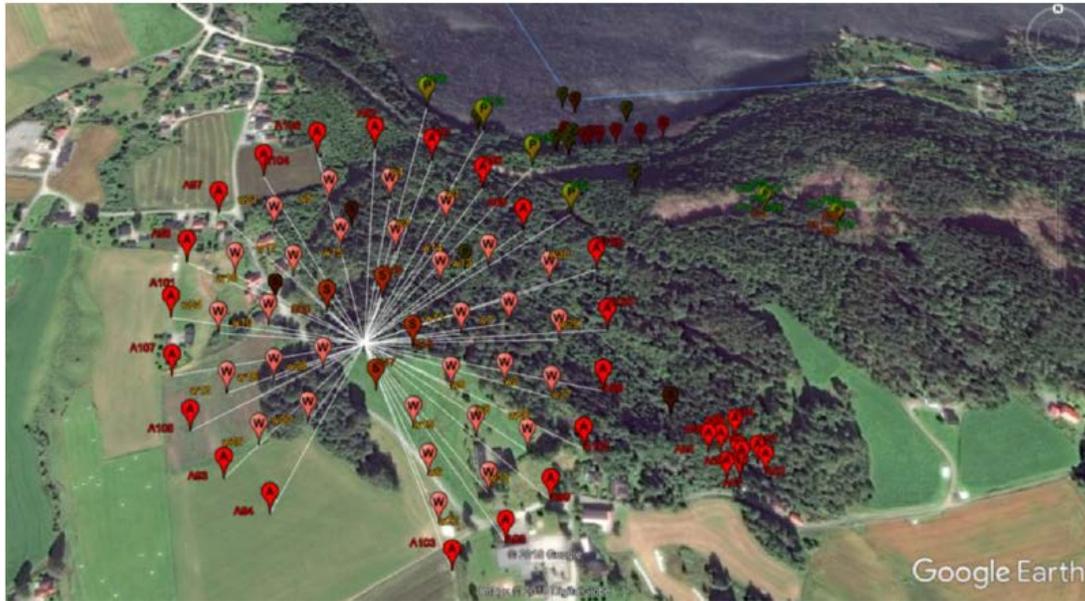


Figure 5.2-1: Sample locations visualised in Google Earth. The red indicators correspond to different sample types.

Finding the fitted distributions and assigning PDFs (for activity concentration data)

As noted earlier, the NORMALYSA software has the option whereby calculations can be run probabilistically. In this way, the uncertainty associated with various input datasets and parameters can be characterised and to some extent accounted for. With this in mind, the distributions associated with some of the datasets we were dealing with (notably the soil-vegetation datasets) were considered in more detail.

Distribution fitting (for activity concentration data)

Standard Probability Density Functions (PDFs) can be fitted to measurement values or samples generated by weighted resampling. The default method of fitting distribution parameters is the maximum likelihood method. After fitting the distribution parameters, the Kolmogorov-Smirnov (KS) test statistics is calculated for each PDF. The KS-statistic is defined as the maximum deviation between the hypothesized cumulative distribution function and the empirical cumulative density function and is a measure of the discrepancy of the tested PDF and the data. The fitted distributions can be ranked in order of decreasing test statistic. Note that the KS test statistic is only one of other possible measures of the goodness-of-fit. It is required that there are at least three observed values to fit distributions to the data.

The default method of fitting distribution parameters is the maximum likelihood method. The values of the parameters of the distribution are then taken as the values that maximize the likelihood function:

$$L(\theta_1, \dots, \theta_k | y) = \prod_{i=1}^n f(y_i | \theta_1, \dots, \theta_k) \tag{5.2}$$

Comparing model results and TLD empirical data for radionuclide activity concentrations in vegetation

Running NORMALYSA probabilistically requires the assignment of PDFs to input parameters such as CRs. For this, the IAEA TRS-479 (IAEA, 2014) has been consulted (Table 5.2-4).

Table 5.2-4: Element specific CR values for various categories of organism from TRS-479 (IAEA, 2014)

Element	CR value, AM±AMSD (vegetation) from TRS-479, kg f.w./kg d.w.	CR value, AM±AMSD (Normalysa vegetation category) kg d.w./kg d.w.
U	2.3E-1 ± 6.4E-1 (shrubs)	2.3 ± 6.4 (understorey, Berries)
U	6.8E-3 ± 1.4E-2 (Tree)	1.36E-02 ± 2.80E-02 (Wood)
U	6.8E-3 ± 1.4E-2 (Tree)	6.8E-2 ± 1.4E-1 (Tree leaves)
Th	2.5E-1 ± 5.6E-1 (shrubs)	2.5 ± 5.6 (understorey, Berries)
Th	1.1E-3 ± 1.1E-3 (Tree)	2.2E-3 ± 2.2E-3 (Wood)
Th	1.1E-3 ± 1.1E-3 (Tree)	1.1E-2 ± 1.1E-2 (Tree leaves)

*Dry weight fraction: Trees (wood) = 0.5; Trees (other parts) = 0.1; shrubs (other parts) = 0.1

The Tables in Mora *et al.* (2018) provide underpinning soil and plant data from Fen complex, extracted from the TLD, grouped and statistically analysed along with the fitted log-normal distributions which have been used as inputs in running NORMALYSA. Additionally, predicted plants activity concentrations made by NORMALYSA are provided in the aforementioned report along with the measured values for comparison. A summary of this information showing predicted versus empirical values is provided in Table 5.2-5.

Table 5.2-5: Summarised information of (NORMALYSA) model-derived and (TLD) empirically-based ²³⁸U and ²³²Th activity concentrations in different components of the Fen forest system.

Sample Group in Milestone 1-3	Plant Type	Th-232 predicted* (Bq/kg)	Th-232 measured (Bq/kg)	U-238 predicted* (Bq/kg)	U-238 measured (Bq/kg)
1	Understorey(a)	9674	24.6	57.5	11.1
	Tree(b)	14.9	0.76	0.5	0.18
2	Understorey(c)	3632	34.7	37	7.3
	Tree (d)	5.2	4.3	0.3	0.2

3	Understorey(e)	2162	28.5	241	5.3
	Tree(f)	3.3	1.5	0.7	0.2
4	Understorey(c)	5876	18.4	122	11
	Tree(d)	9.6	5.7	1	0.5
5	Understorey(g)	4037	50.8	47	2.9

*Median value from probabilistic simulations

Empirical data (a) = fern and moss; (b)Picea abies; (c) moss (d) pine and Picea abies; (e) moss and lichen; (f) birch and Picea abies; (g) fern (plant&root)

The correspondence of modelled activity concentrations with concomitant field-based activity concentrations in vegetation are generally quite poor. In some cases, such as those related to ²³²Th activity concentration in understorey vegetation, the model provides an over-prediction that can be greater than a factor of 100. The closest correspondence between model output and measured values was observed for ²³⁸U in trees (Figure 5.2-2), but the number of samples involved in the analysis (n=4) renders it impracticable to draw any firm conclusions concerning the efficacy of the model.

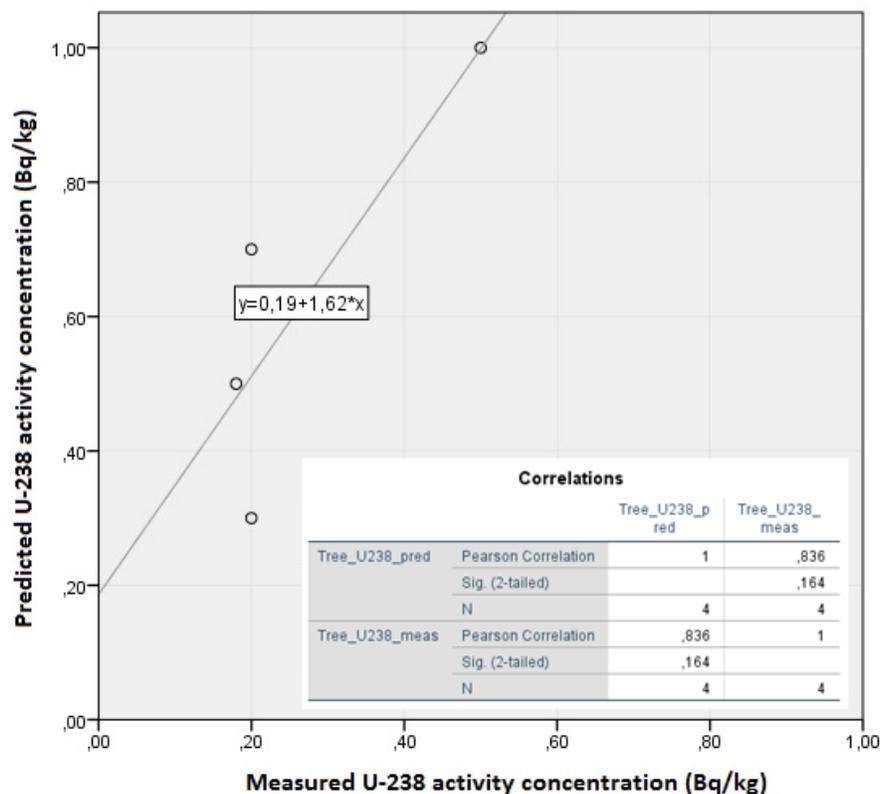


Figure 5.2-2: Comparison of model-predicted and measured activity concentrations of ²³⁸U in trees.

There are a few issues to have in mind when we do this kind of comparison. By way of example, the following two points are quite relevant:

- Inhomogeneity of the contamination levels in the Fen complex mean that is important that soil and plant samples are taken from exactly the same locations. However, this is not always the case and errors may have arisen because of this discrepancy.
- Plant categories in NORMALYSA are generic. We have to map specifically measured samples onto generic categories and this undoubtedly introduces uncertainty.

In general, comparing the outputs of NORMALYSA with empirical data (on activity concentrations) provides very few insights from which definitive conclusions can be drawn. However, some general observations can be made as will be addressed at the end of this section, once modelling of external exposures have been considered.

Comparing model results and TLD empirical data for ambient dose equivalent data

In the 'Assessment Context' set-up for NORMALYSA, radionuclides within the ^{232}Th and ^{238}U decay series were selected and a simple 2-component model configured, comprising of waste material overlain by a cover layer. The cover thickness was set to 0 m for the sake of the model- empirical data comparison. For our purposes, it was most relevant to take predictions at the start of the simulation period as we were not attempting to make forecasts about the evolution of contamination levels (and exposure levels). Probabilistic data provided the most useful outputs (Figure 5.2-3).

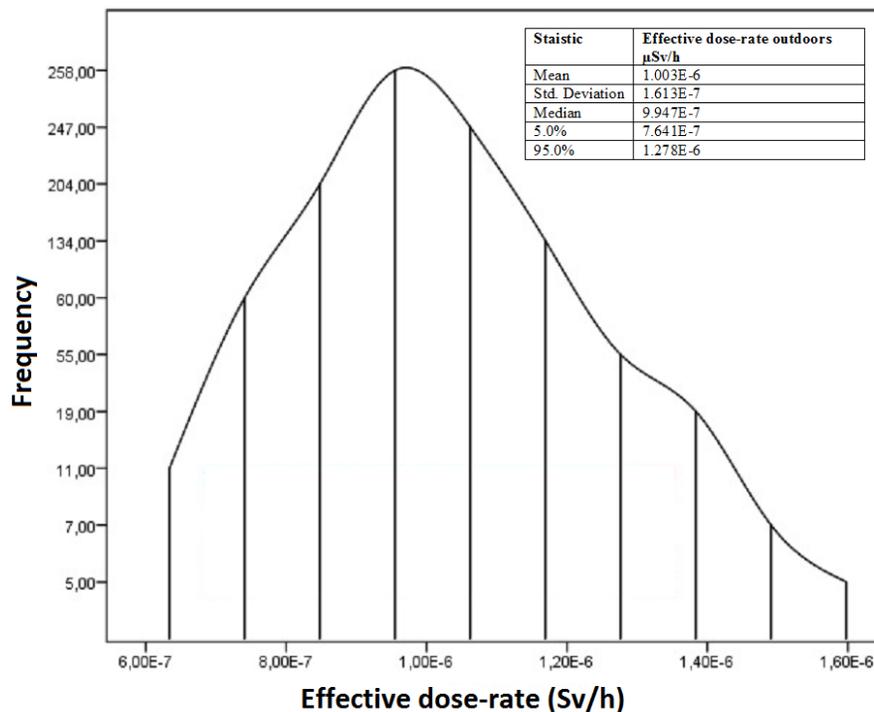


Figure 5.2-3: Probabilistic output from NORMALYSA for selected sites (S1 to S26: see the TLD) at Fen.

A direct comparison between model outputs and empirical data is not possible without some consideration of various factors that confound this aim, notably

- A value of $240 \mu\text{Sv y}^{-1}$ ($0.027 \mu\text{Gy h}^{-1}$), derived from UNSCEAR (UNSCEAR, 2000) has been used to account for the cosmic component of radiation in the calculations. This value is for sea-level but the error introduced would be minimal considering the fact that the Fen field sampling area lies at an elevation of approximately 200 m above sea level.
- NORMALYSA provides output in terms of ambient dose rate equivalents. A conversion is, therefore, required to provide output in the same units as the measurements data- provided as gamma air kerma rates. In the energy range 0.3 to 3 MeV the ratio $H^*(10)/K_a$ (Sv Gy^{-1}), lies in the range from 1.31 to 1.13 (ICRP, 1996). For sake of simplicity, a conversion (corresponding to 1.5 MeV) of 1.15 has been used in the subsequent calculations.

Secular equilibrium has been assumed for the considered decay chains and a porosity = 0.2 was applied based on empirical data (there was a requirement to change the NORMALYSA default, to match other model, i.e. CROM and GRANIS, applications). Furthermore, the contribution from ^{40}K was derived by establishing an approximate activity concentration for this radionuclide in soil from empirical determinations (as collated in the TLD). An indicative value of 500 Bq kg^{-1} was assumed in line with the previous analysis (Mora *et al.*, 2018). The model was run probabilistically – since at some sites multiple samples of soil had been taken and determinations of decay chain radionuclides made it practicable to define an arithmetic mean and standard deviation in some cases. A lognormal distribution was thereafter assumed. When models were run probabilistically, the arithmetic mean from the output was used as the best estimate model prediction for the given site. The results of the analysis are provided in Figure 5.2-4.

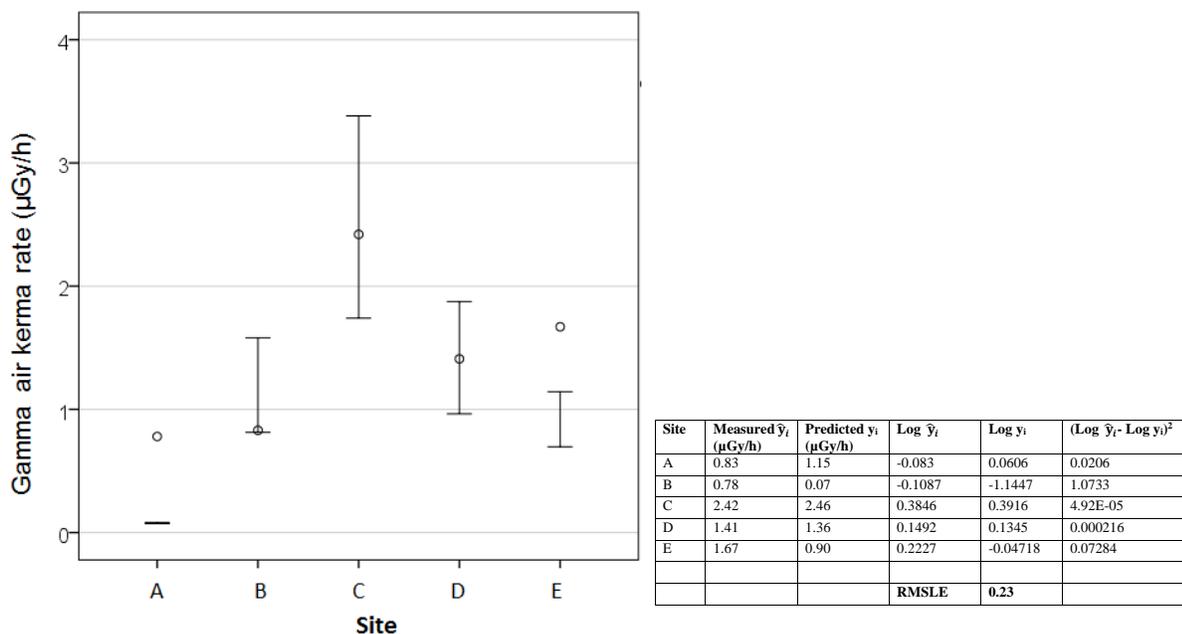


Figure 5.2-4: Measured and predicted Gamma air kerma rates ($\mu\text{Gy h}^{-1}$) for 5 sites at Fen. Error bars show 5th and 95th prediction for model output and the circle provide the measured data. For sites C and D, modelled values are indicative, the percentiles were estimate from the most extreme case where a range in values was expressed (Site B)

Comparisons between modelled and empirical data using the guidance provided in Section 3. The root mean squared logarithmic error (RMSLE) was used to provide a measure of model performance (as in Equation 3.17).

A RMSLE of 0.23 was calculated for NORMALYSA based on a comparisons of 5 coupled predicted (model output) and measured data points. Other ‘performance’ metrics have also been derived as shown later in this report.

Some general observations can be made from the application of NORMALYSA at the Fen site. It is important to establish how coupled the model and input datasets are and establish who the user is (a Developer? Experimentalist or both?). Ideally, one would have access to a closely integrated group (good communication between modellers and experimentalists) so that input datasets are tailored to the model and the model output is closely linked to the experimental data you are collating (if you want to test a model for example). For the Fen NORMALYSA application, the limitations are shown when the modeller and experimentalist are not closely linked. It is not possible to say whether NORMALYSA is fit-for-purpose at Fen nor whether further process based modelling would be appropriate.

It is clearly important to have comprehensive documentation for model testing – we need to understand the details of the model (e.g. see underlying equations etc.) to make sure that the model is configured/parameterised correctly. The question regarding how many data are required is always challenging– we never seem to have an answer where ‘n’ must be greater than a certain number because it is always case dependent. Having the “right” data may suggest there being compatibility with the requirements of the model (measured value corresponds directly to required model parameter)

It seems apparent that the more you need to interpret/ torture data the more uncertainty seems to be introduced. Furthermore, care should be taken over (getting the right) conversion factors and units when applying the model. The best fitted PDF should be physically reasonable, i.e. they should not violate the physical laws of nature. It seems clear that by making sweeping generalisations, there is a corresponding increase in “uncertainty”. Sophistication, in the form of more process-based modelling also incorporates limitations (parametrisation and site specificity). It seems that the “right” balance needs to be found between a model that is simple enough to parameterise and explain (to stakeholders) and nonetheless provides robustness in terms of being able to simulate the key processes in a convincing manner.

5.2.3 GRANIS

The GRANIS model has been used to calculate air kerma one metre above the ground using measured activity concentrations of NORM in different depths of soil. The measurements have been made in an area of fenland in the Telemark region of Norway. The geology of the region comprises magmatic carbonatite rocks and is known to be a large natural reservoir of NORM in particular thorium (^{232}Th). As a consequence of historical mining activities the soils in this area are contaminated with deposits of ^{238}U , ^{232}Th and their progeny. The composition and density of the soil is used in the GRANIS model to generate material specific attenuation coefficients. It is recognised that soil composition is much less important than soil density in the calculation of external exposure for the photon energies of interest in environmental assessments. Nevertheless, a representative soil composition for the fen site has been derived using data supplied by NRPA and used in this assessment (Table 5.2-6). A soil density of $1.6 \cdot 10^{-3} \text{ kg cm}^{-3}$ wet weight has been assumed and with a porosity of 20% a soil density of $1.4 \cdot 10^{-3} \text{ kg cm}^{-3}$ wet weight has been calculated.

Activity concentrations in soil have been converted to activity concentrations per unit volume for input to GRANIS based on these densities.

Table 5.2-6: Assumed composition of fenland soil

Element	Fraction of mass	Element	Fraction of mass
Al	3.22E-02	Mg	1.29E-02
Ba	2.36E-03	Mn	1.93E-03
Pb	1.18E-04	Na	1.18E-02
Ce	4.73E-04	S	1.72E-03
P	1.07E-02	Ti	3.65E-03
Fe	1.61E-01	H	2.22E-02
Ca	1.18E-01	C	2.00E-01
K	1.72E-02	O	1.78E-01
Si	2.26E-01		

Some of the soil concentration measurements are co-located with air kerma measurements and can be used to assess the performance of the model. In all there are five locations where soil concentrations and air kerma measurements are co-located and these are reported in Table 5.2-7 and Figure 5.2-5 along with the results of the GRANIS model. The locations A to E are the same as those considered in the NORMALYSA assessment. Some slight differences in the measured air kerma rates are due to adjustments made to remove the contributions from ^{40}K and cosmic radiation.

The comparison of model predictions and observations shows that differences vary by factors of between 1.2 and 42. This is in part due to the extreme heterogeneity of the contamination and hence the sensitivity of measurements to the precise location of sampling. For example, at the location (59.28165, 9.2861) the measured ^{232}Th soil concentration is $6.8 \cdot 10^1 \text{ Bq kg}^{-1}$ and the measured air kerma is $7.3 \cdot 10^{-1} \mu\text{Gy y}^{-1}$ however at (59.27708, 9.303766) for the same sample depth the corresponding measurements are $2.1 \cdot 10^3 \text{ Bq kg}^{-1}$ and $1.6 \cdot 10^0 \mu\text{Gy y}^{-1}$, which appears to be inconsistent.

A value of 0.77 has been calculated for the root mean squared logarithmic error (RMSLE) which can be compared with the RMSLE for other models that have been applied to this data. It can be seen that for this dataset that the RMSLE value calculated using NORMALYSA is lower than that for GRANIS and hence is a better fit to the observations. The NORMALYSA model would appear to be based on a simpler calculation methodology than GRANIS and therefore the results would appear to contradict the assumption that processed based models are necessarily better than simpler ones. However, it is possible that simpler models will give reasonable predictions for the wrong reasons such as cancellation of errors or inaccurate representation of the scenario being modelled. The latter seems counter intuitive but can

be explained by considering the application of NORMALYSA to this dataset where soil depths at some locations were just a few centimetres but were modelled assuming uniform contamination over a much greater depth ie a soil depth of 3 m. This has the effect of increasing predictions of gamma air kerma rates at these locations and because some of the observed soil concentrations appear to be lower than expected a better match with observations is achieved. It can be seen from Table 5.2-7 and Figure 5.2-5 that it is these locations (Sites A and E) where GRANIS performs most poorly. Other statistical comparisons can be found in the Appendices.

Table 5.2-7: Comparison of predicted and observed air kerma rates at 5 Norwegian Fen sites

Depth	Latitude	Longitude	Radionuclide	Soil Conc FW (Bq kg ⁻¹)	Soil Conc DW (Bq kg ⁻¹)	GRANIS Air Kerma (uGy h ⁻¹)			Measured Air Kerma (uGy h ⁻¹)*		
						Mean	-1 stddev	+1 stddev	Mean	-1 stddev	+1 stddev
0to0.05m A	59.28165	9.2861	U-238		3.82E+01	1.74E-02	1.53E-02	1.95E-02	7.32E-01	4.10E-02	1.42E+00
			Th-232		6.79E+01						
0to0.10m E	59.27708	9.303766	U-238		1.29E+02	5.09E-01	5.03E-01	5.16E-01	1.62E+00	1.08E+00	2.16E+00
			Th-232		2.33E+03						
0to0.25m ~ C	59.27427	9.300366	U-238	4.70E+02	5.37E+02	1.90E+00	1.81E+00	2.00E+00	2.38E+00	1.05E+00	3.70E+00
			Ra-226	9.60E+01	1.10E+02						
	Th-232	6.24E+03	7.13E+03								
	U-238	4.10E+02	4.69E+02								
D	59.27702	9.30405	Ra-226	9.60E+01	1.10E+02	1.05E+00	9.97E-01	1.10E+00	1.37E+00	7.78E-01	1.96E+00
			Th-232	3.38E+03	3.86E+03						
0to0.05m 0.05to0.1m 0.1to0.2m B	59.2819	9.286417	U-238		1.87E+03	1.31E+00	1.04E+00	1.59E+00	7.79E-01	-1.31E-04	1.56E+00
					1.46E+03						
					1.01E+03						
			Th-232		2.60E+03						
					2.02E+03						
					1.12E+03						

* Contributions to air kerma from ⁴⁰K and cosmic radiation have been removed.

~ For these GRANIS results uncertainties are based on measurement uncertainties in soil concentrations, all others are based on variability in soil concentration measurements.

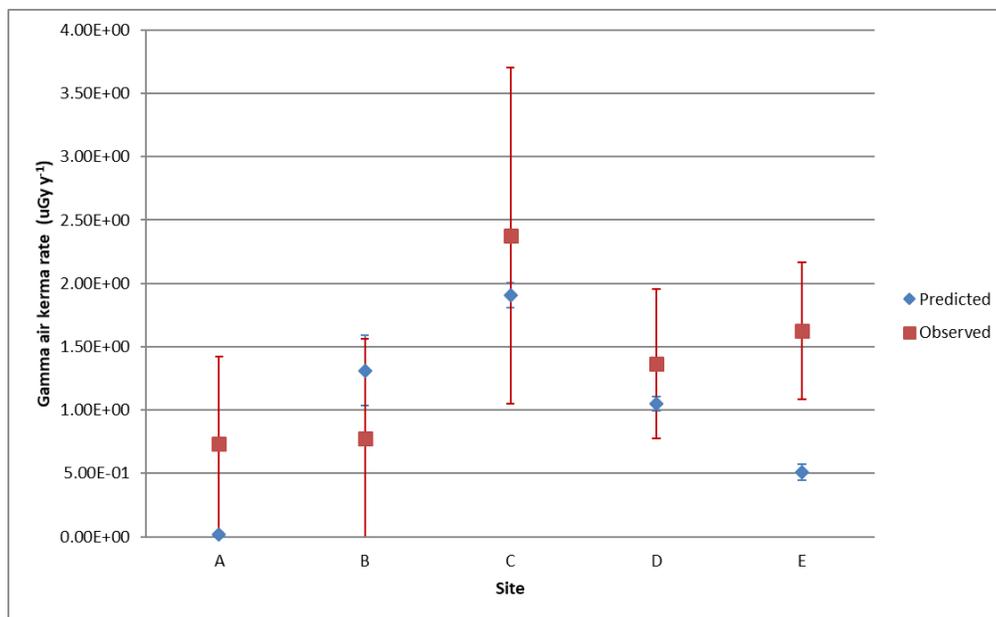


Figure 5.2-5: Comparison of predicted and observed gamma air kerma rates ($\mu\text{Gy h}^{-1}$) for 5 sites at Fen. Error bars show one standard deviation for model output and measured data.

6 CASE STUDY: BELGIAN NORM SITE

6.1 Introduction to the case study

The Belgian NORM site (Tessenderlo site in Belgium that is contaminated by NORM industry discharges) is a calcium difluoride sludge heap from the phosphate industry partly covered with vegetation such as pine trees, birch trees, grasses, shrubs, etc. The sludge heap was used until 1979 for the deposition of CaF_2 -sludge coming from the phosphate industry. Currently, the heap is partly covered with vegetation and the north-eastern part is already being used as waste storage from clean-up/decontamination activities. The site can thus be subdivided into a calcium difluoride sludge heap and a holding pond for wastewater. The observatory site is situated on the sludge heap.

The site chosen for investigation as part of the present project is mostly a forest of pine trees, mainly *Pinus sylvestris*, and the dimensions of the area are about 200 by 200 m. Besides pine, the following (mainly native) vegetation can be distinguished: *Betula pendula*, *Salix caprea*, *Quercus rubra*, *Eupatorium cannabinum*, *Tanacetum vulgare*, *Tussilago farfara*, *Epilobium hirsutum*, *Jacobaea vulgaris*, *Cirsium arvense*, *Cirsium vulgare*, *Sonchus sp.*, *Artemisia vulgaris*, *Chenopodium sp.*, *Epipactis helleborine*, *Humulus lupulus*, *Rubus sp.*, *Phragmites australis* and *Calamagrostis epigejos*. The site has a temperate maritime climate, characterised by cool/humid summers and mild/rainy winters.

Radionuclides present at the site, mainly ^{238}U and its progeny, originated from the phosphate ores. Contamination levels of ^{226}Ra between 2500 and 3500 Bq kg^{-1} can be found in the soil and sludge together with co-contaminants such as As, Cd, Cr, Pb and Zn. Although remediation measures are planned, approximately 7 ha of this site are available for the next 10 to 15 years to perform long-term radioecological research in a NORM-contaminated terrestrial ecosystem.

The ambient gamma-dose rate in the contaminated area is 400 – 800 nSv/h, compared to about 100 nSv/h, as background in the surrounding area. Two preliminary measurement campaigns have been provided there, paying the main attention to the soil and pine trees, but also analysing the grass and moss:

- October 13, 2016: a pine tree, grass (only Pb-210 above detection limit in tree samples, also Th-228 and Ra-226 in grass).
- November 28, 2017: pine trees, grass, moss (3 samples each, Pb-210, U-238, Ra-226).

In the course of TERRITORIES, the site has been subject to systematic sampling and monitoring, by installation of an instrumented forest station and seasonal sampling of soil and vegetation with focus on U-238, Ra-226 and Pb-210. As part of this, four seasonal sampling campaigns were performed, between the end of 2017 and the end of 2018. In Addition, the site has been instrumented and being continuously monitoring for rainfall, solar irradiation, litterfall, sap flow of the trees and soil moisture.

6.2 Application of the models to case study

6.2.1 FORESTCROM

The case of Tessenderlo is a NORM case, where trees are growing above a pond of CaF₂ waste enriched in Ra-226. Measurements on different parts of the trees are gradually incorporated in the TLD, for different radionuclides. The final measurements obtained in the site (May 2019) are reflected in the TLD and also in the graphs in this chapter.

A total tree approach using FORESTCROM was applied, calculating the weighted average concentration in the whole tree by weighting the measurements in roots (19.1%), stem and bark (33.9%), branches (25%) and needles (22%). The resulting concentration ratios from the final results and comparative values provided below.

Table 6.2-1: Weighted activity concentrations of Pb-210, U-238 and Ra-226 measured in pine trees and soil, campaign: May, 2019, and concentration ratios derived experimentally and from CRs in the Erica tool and CROM databases for trees coniferous.

Fraction	Pb-210	U-238	Ra-226
	Concentration, Bq/kg	Concentration, Bq/kg	Concentration, Bq/kg
Total tree	79.5	108.2	120.5
Soil (bottom)	3076.5	6102.9	4438.8
	Concentration ratio	Concentration ratio	Concentration ratio
CR measured	0.032	0.019	0.031
CR from TRS 479 (GM)	0.0697	0.0066	0.0116

The values provided in Table 6.2.-2 were used for the model. The physical decay constant for Pb-210 was considered to be equal to the constant for Ra-226, as secular equilibrium was assumed in the decay chain.

Table 6.2-2: Values used for of the application of FORESTCROM in Tessenderlo forest – Pb-210.

NAME	VALUE	REFERENCE
λ_{ec-t} (d ⁻¹)	6.5E-03	From IAEA TRS472 (IAEA,2010) - half lives from 3-4 weeks to 3 months in growing phase, to 4-6 months in dormancy phase.
CR	7.0E-02	From Erica-tool (v. 2015) or CROM (v. 8.3.0)
ρ (kg/m ³)	1.3E+03	Usual soil density from 1300 to 1500 (lower value)
h (m)	2.0E-01	20 cm depth for the roots is assumed
λ_{soil} (d ⁻¹)	0	From IAEA SRS19 (IAEA, 2001) pg 65.
λ_{phy} Pb-210 (d ⁻¹)	1.2E-06	Decay constant for Ra-226, as Pb-210 is considered to be in secular equilibrium with that isotope

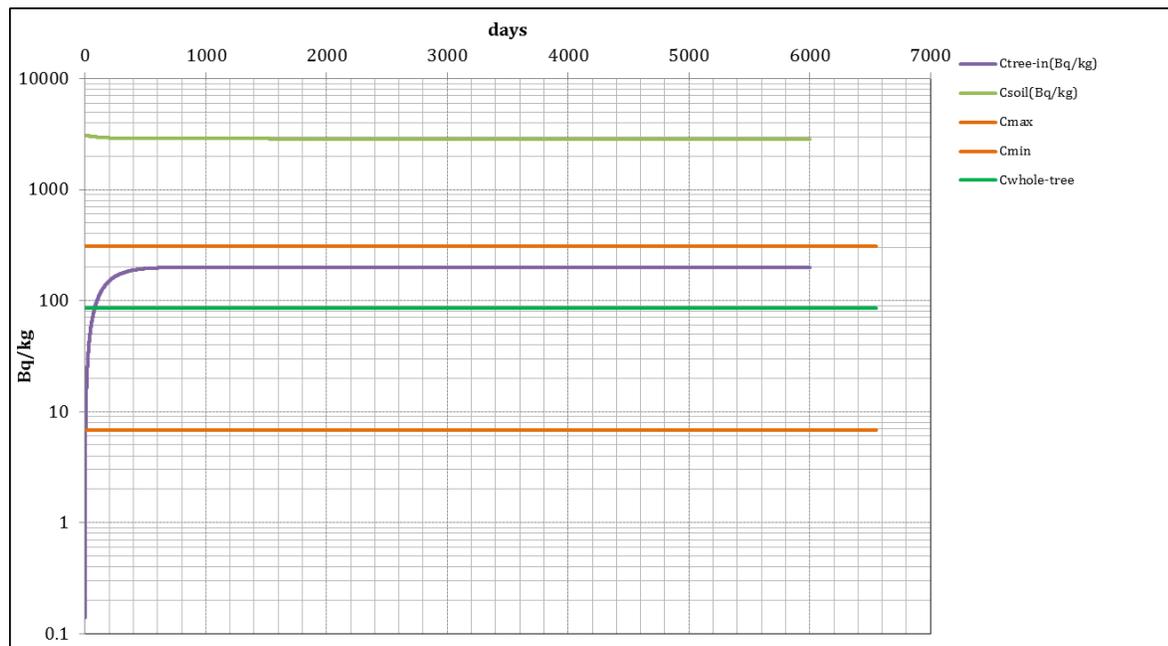


Figure 6.2-1: Results of the model using the parameters from Tables 6.2-1 and 6.2-2 for Pb210 (obtained without information from Tessenderlo) and comparison with the measurements in the tree as a whole.

The FORESTCROM model, compared with the minimum and maximum values measured for Pb-210 in Tessenderlo in different parts of the pine trees, for the whole pine trees, using parameters from

bibliography provides a good agreement, being conservative with respect with the final measurements. In the figure minimum and maximum obtained concentrations for the different parts of the tree are represented. The agreement with measurements in the pine tree as a whole is good in the equilibrium.

Comparisons of the model with other measured radionuclides were performed (Table 6.2-3 and Figure 6.2-2 for Ra-226 and Table 6.2-4 and Figure 6.2-3 for U-238).

Table 6.2-3: Values used for FORESTCROM model in Tessengerlo forest – Ra-226.

NAME	VALUE	REFERENCE
λ_{ec-t} (d ⁻¹)	6.5E-03	From IAEA TRS472 (IAEA, 2010) - half-lives from 3-4 weeks to 3 months in growing phase, to 4-6 months in dormancy phase.
CR	1.2E-02	From Erica-tool (v. 2015) or CROM (v. 8.3.0)
ρ (kg/m ³)	1.3E+03	Usual soil density from 1300 to 1500 (lower value)
h (m)	2.0E-01	20 cm depth for the roots is assumed also
λ_{soil} (d ⁻¹)	0	From IAEA SRS19 (IAEA, 2001) pg 65.
λ_{phy} Ra-226 (d ⁻¹)	1.2E-06	Decay constant for Ra-226

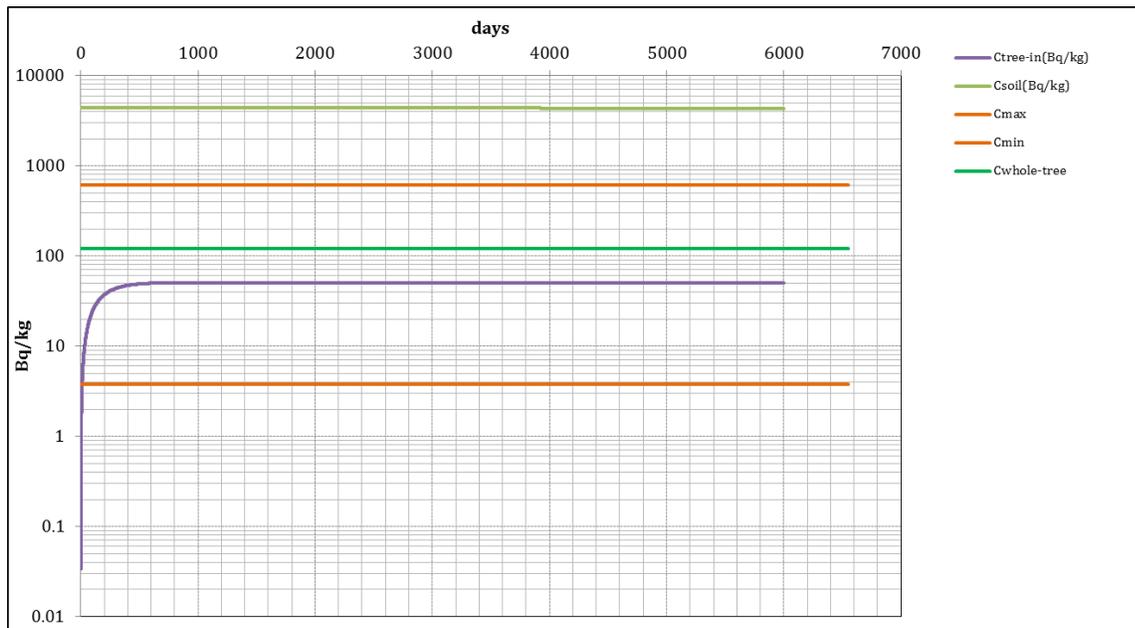


Figure 6.2-2: Results of the model using the parameters from bibliography for Ra-226.

Table 6.2-4: Values used for FORESTCROM model in Tessengerlo forest for U-238

NAME	VALUE	REFERENCE
λ_{ec-t} (d ⁻¹)	6.5E-03	From IAEA TRS472 (IAEA, 2010) - half-lives from 3-4 weeks to 3 months in growing phase, to 4-6 months in dormancy phase.
CR	6.6E-03	From Erica-tool (v. 2015) or CROM (v. 8.3.0)
ρ (kg/m ³)	1.3E+03	Usual soil density from 1300 to 1500 (lower value)
h (m)	2.0E-01	20 cm depth for the roots is assumed
λ_{soil} (d ⁻¹)	0	From IAEA SRS19 (IAEA, 2001) pg 65.
λ_{phy} U-238 (d ⁻¹)	4.3E-13	Decay constant for U-238

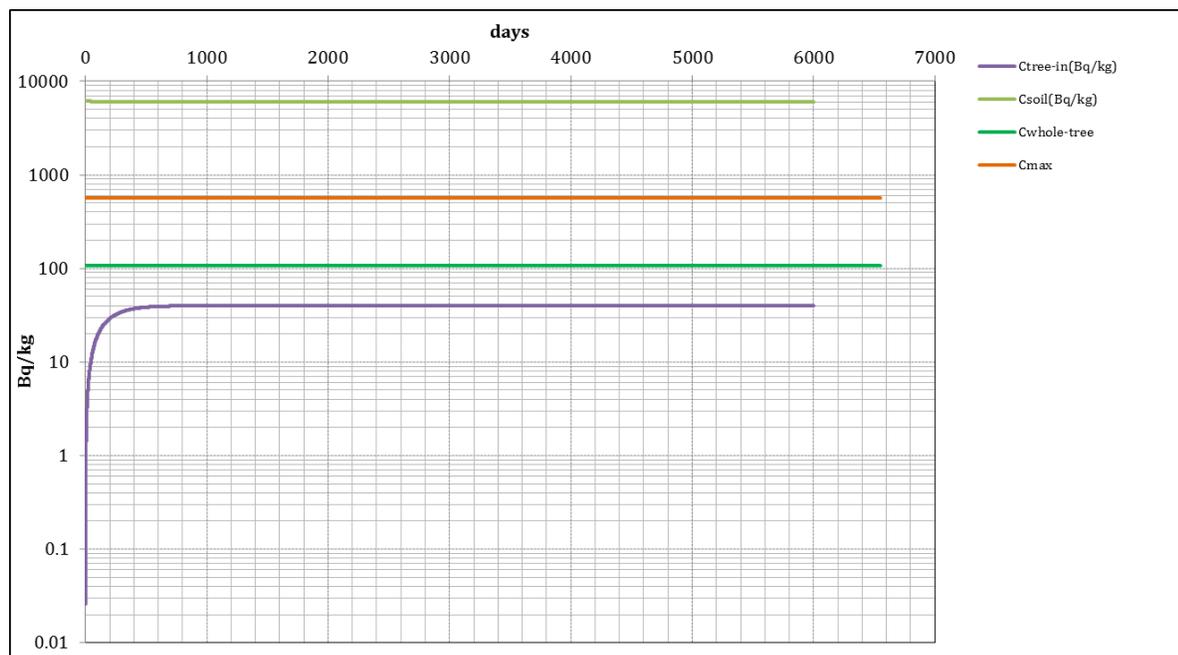


Figure 6.2-3: Results of the model for U-238.

In both cases, for U-238 and Ra-226 the model underestimates the weighted value for the whole tree. Obviously this is caused by the CR considered for the radionuclides, as they correspond to normal soils, while in this case the material where the trees are growing has very different properties. If CRs are calculated for this material (0.0336, 0.0207 and 0.035 for Pb-210, U-238 and Ra-226 respectively - information provided by Jordi Vives-i-Batlle - 28th of May of 2019), the results in the figure 6.2-4 are obtained.

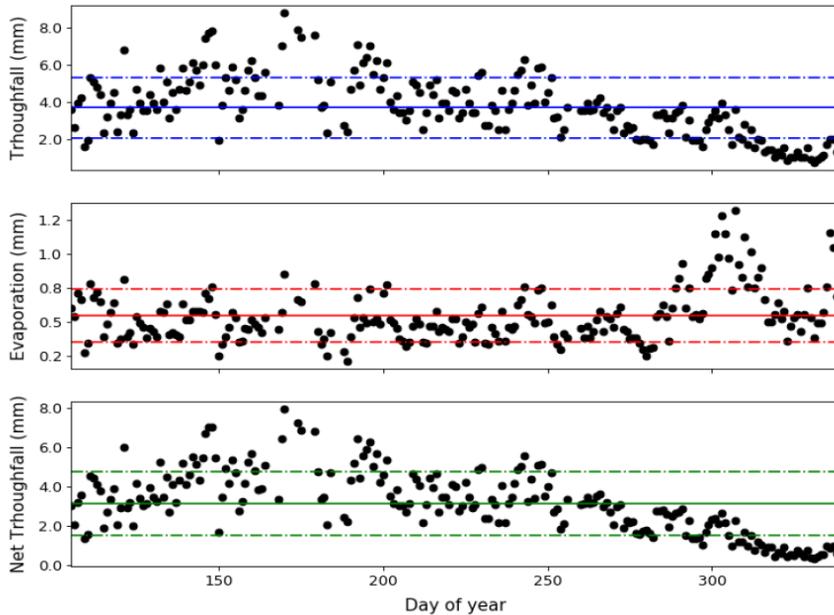


Figure 6.2-5: Weather data over a typical year

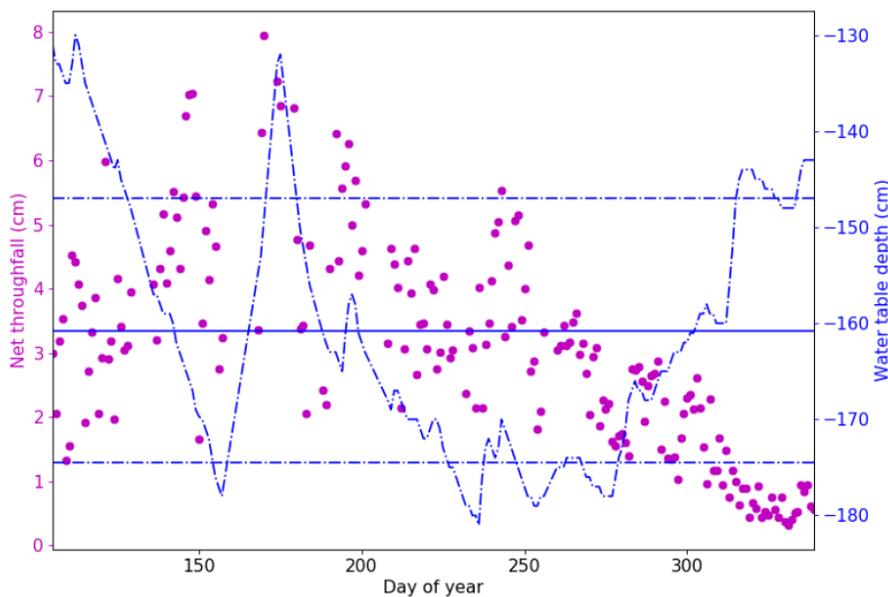


Figure 6.2-6: Water table fluctuations in a typical year in response to atmospheric conditions, as measured by soil moisture probes

The K_d is a key parameter in the forest models, regulating as it does the transport of radionuclides in the soil. The K_d controls the retardation and the exchange between available and the unavailable fractions of the radionuclides. Experimental determination of the K_d in situ in this project was fraught with difficulties, because the sludge (CaF_2) was found to be so compact that we could not mechanically press-out water and concentrations in the interstitial water were persistently below detection limit. The K_d is therefore best obtained from laboratory experiments. Since it is difficult to obtain reproducible K_d s in laboratory based on column methods, we settled for a batch-extraction

method in the laboratory. We determined K_d s for uranium in the laboratory, using a pH = 5 buffered solution to simulate plant root acidification, in time for this report (see results below). At the time of writing, K_d experiments are available for radium (using Ba as an analogue), polonium and lead.

Implementation of ECOFOR to the Belgian NORM site

The first step in implementing ECOFOR to the NORM site was to optimise the hydrological module, then test it for the Belgian Mol forest and then re-parameterise the model for the Belgian NORM site. The reason for the approach is that there is an abundance of hydrological validation data for the Mol forest. The Mol forest (sandy soil), has a high permeability – high hydraulic conductivity, heterogeneous 6-layer horizon, low field capacity (FC) and free drainage water circulating with moisture content for each layer being at around FC. The Belgian NORM site (Phosphate CaF_2 sludge), conversely, is nearly impermeable because it has a low hydraulic conductivity, simpler structure with one upper organic layer followed by a homogeneous CaF_2 profile, high field capacity, slow drainage with moisture content for each layer at around porosity.

Once we verified the best solution to reproduce the complex time series of hydrological data (moisture content in different soil layers) we acquired sufficient confidence that the algorithm is correct, and by re-parameterising it for the CaF_2 sludge typical of this site, we obtained a new version of the model for the radionuclides ^{238}U , ^{210}Pb and ^{226}Ra .

Model improvements

During the improvement of the hydrological model, particular attention was paid to balancing hygroscopic, capillary and gravitational water: from an initial application of Richards equation, we considered a range of process simplifications ranging from the different versions of the equation (low flow velocities, relatively saturated conditions) down to the simpler ‘tipping bucket’ approach, reaching an intermediary compromise in which tipping bucket switches apply but the water velocities are not instantaneous but defined by the Lucas Washburn equation for upward capillary motion, and the Darcy law with a variable rather than saturated hydraulic conductivity for downward motion. We tried different equations for a more realistic, soil moisture-dependent formulation of the hydraulic conductivity ranging from the Van Genuchten (1980), Brooks and Corey (1966), Campbell (1974) and Kendy *et al.* (2003) formulas, settling for the Kendy exponential equation for ease of computation and minimisation of parameters. The selected formula was validated by running soil water profiles in the Mol forest (for the case of free water drainage through the soil column), where more hydrological data was available to compare.

Further model modifications include the fact that “tipping bucket” switches act automatically on every donor or acceptor compartment which are open only when soil moisture in a given soil layer is between the residual and saturated water content. Also worthy of note is a simplified representation of elements’ linkage to water was adopted through retardation processes in soil (rather than using the complex coupling between transport and advection/diffusion equations).

ECOFOR parameterisation for the Belgian NORM site

To implement the ECOFOR hydrological sub-model, we devised as soil profile with 9 layers of constant thickness (1 m), knowing that the CaF₂ sludge is uniform. On top of these layers there is a surface 20-cm layer of “normal” topsoil. We used hydrological parameters for the sludge from (Bartos and Palermo 1977) (Table 6.2-5). However, we are presently in the process of replacing these parameters with direct measurements on site, as more is known about the texture of the sludge.

Table 6.2-5: Main characteristics of CaF₂ sludge as previously reported in the literature

Parameter	Value	Units	Source
Median grain size	2.00E-05	m	Bartos and Palermo, 1977
Bulk density	2.76E+03	kg m ⁻³	Bartos and Palermo, 1977
Particle density	8.90E+03	kg m ⁻³	Bartos and Palermo, 1977
Saturated Hc	2.75E-02	m d ⁻¹	Bartos and Palermo, 1977
% Al₂O₃ =	1.8625		Vanhoudt, 2015
%SiO₂ =	27.74		Vanhoudt, 2015
% clay =	5.59%		Assumes clay has 1 part Al ₂ O ₃ and 2 parts SiO ₂
% sand =	24.02%		Assumes clay has 1 part Al ₂ O ₃ and 2 parts SiO ₂

The most important and uncertain hydrological parameters are the field capacity (FC) and hydraulic conductivity (Hc), which determine soil water hold-up capacity of the soil and the water flows upon which the rest of the model entirely depends. For the transport of elements in soil, the most uncertain parameter is the distribution coefficient (K_d), controlling retardation. The K_d is the main parameter in our model, regulating as it does the transport of radionuclides in the soil. It also controls the retardation and the exchange between available and the unavailable fractions of the radionuclides.

For the field capacity in sludge we used the general indication that it is about 15 to 25% for sandy soils, 35 to 45% for loam soils, and 45 to 55% for clay soils (<https://nrcca.cals.cornell.edu/soil/CA2/CA0212.1-3.php>). Sludge would be more akin to clay in that it is a fine impermeable material – So we use a FC = 0.5. The saturated hydraulic conductivity was derived from Bartos and Palermo (1977) and the moisture-dependent (unsaturated) Hc is automatically updated at each time step by the model as a function of the moisture content, based on the Keady approximation (Keady *et al.*, 2003).

Another important consideration in the model was the exchange rates between available and the unavailable fractions of the radionuclides in soil i.e. the unavailable and available soil fractions of the linear reversible model. They can be expressed as a function of the K_d and the effective diffusion rate K_A between unavailable and available soil fractions of the aforementioned linear reversible model. The model has a large uncertainty in the diffusion rate, but in practice this does not affect significantly the model output because vertical transport of radionuclides is much slower than sorption and in this condition the sorption

rate is not a very sensitive parameter and, after several calibration trials, an optimisation value of 0.015 d^{-1} was chosen, consistent with a representative value for Sr in loam soils $k_{14} = 0.01 - 0.03 \text{ d}^{-1}$ (Vandenhove *et al.*, 2009) and this was used for all elements until more detailed data arrives. Finally, we had to estimate residual water content (this is presently judged to be at 2%), the wilting point, estimated to be 20% (typical of clay) and the saturation point, calculated using the soil texture triangle (<http://resources.hwb.wales.gov.uk/VTC/env-sci/module2/soils/soilwatr.htm>) which gives 55.5%.

For the vegetation ET part of the model, the most important improvement was to adjust plant water transport based on having root uptake balanced against evapotranspiration (water mass conservation). The evapotranspiration model was upgraded from a simpler version in which monthly solar irradiation data is considered, to one in which we used daily time steps using the more precise solar irradiation data for the correct longitude and latitude of the site, as given by the Helioclim database (<http://www.soda-pro.com/web-services#top>). Figure 6.2-7 shows the improvement by comparing model output for these two cases (red) with our field data (blue).

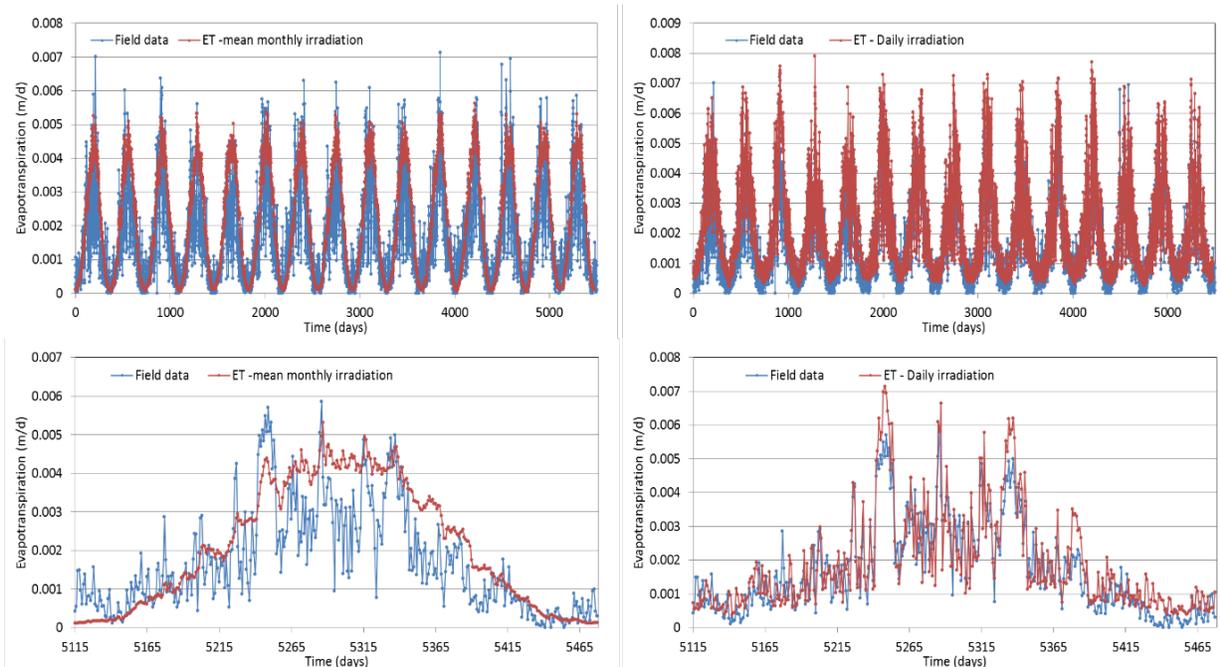


Figure 6.2-7: Improvements in the evapotranspiration model of ECOFOR

Another improvement was to derive empirically selectivity coefficients (SC) for stable elements in pine trees from data for the nearby Mol site (Gielen, Vives i Batlle *et al.* 2016) (Figure 6.2-8) and use this information to link element fluxes to water fluxes in plants, as determined by an independent model designed to do this calibration. Using SCs for trees from the Mol site is in principle justified being the same species and climate than in the NORM site, but health status of the NORM site's trees is different and this could affect the uptake rates, signalling the direction of future investigations.

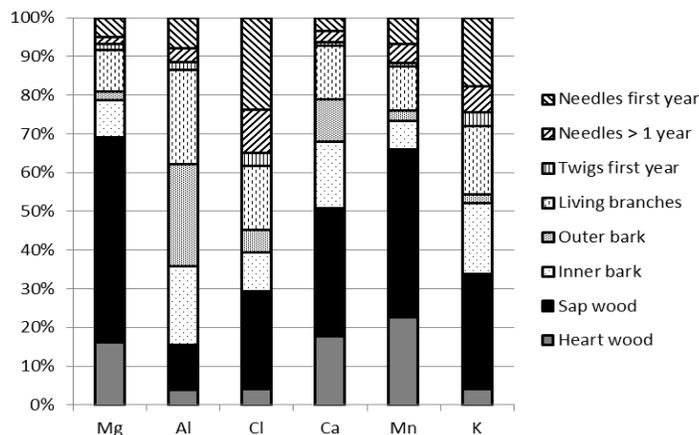


Figure 6.2-8: Percentage of stable elements in different tree parts (Gielen *et al.*, 2016)

We used a calibration procedure as follows: (a) we estimated the SC of ^{238}U , conjecturally, as an average of the essential nutrients K, Mg and Ca (average = 1.7×10^{-2}). Then, for ^{226}Ra and ^{210}Pb , we scaled this value as per ratios with transfer factors of these elements relative to uranium, as found in a previous study (Vandenhove *et al.*, 2009). The values were surprisingly adequate to approximate the distributions of radionuclides observed in pine trees from the Belgian NORM site, and with additional parameter optimisation with U, Ra and Pb-specific data from the site, we optimised these coefficients (usually within a factor of 2 or 3) and now the model provides a very good representation of the data.

ECOFOR results

ECOFOR had been previously validated for the MOL site, and in this context it was found to successfully predict the daily potential evapotranspiration results for pine trees based on Mol climatological data for 1984-1998, and compatible predictions such as: mean soil water content $\sim 25\%$; water in roots > rest of plant > leaves; water table 2.1 – 2.6 m; transpiration $< 1.2 \text{ mm d}^{-1}$ ($\sim 15\%$ of PET at peak times), max. capillarity flow of 0.4 mm d^{-1} . The solution for ^{36}Cl for a constant water table was consistent with the predictions of a previous stand-alone ^{36}Cl model. It also matched the distribution of stable elements (Mg, Cl, Ca, Mn and K) because the model was calibrated with experimentally selectivity coefficients specific for the site. When we compared model output with the hydrological data for the Mol station for the various layers, the predicted time evolution of θ approximates the 2005 data available at SCK-CEN from where this was experimentally determined. This approximation improved further when fine-tuning the saturated hydraulic conductivity in the model.

The model is designed to be transferable, so we could readily apply it for the NORM site with the same processes, but a different set of parameters as previously described. In this way, the hydrological calculations show an important difference, which although it cannot be validated due to the recent installation of the soil moisture sensors in the NORM site, does coincide fully with our initial observations, and our understanding of the hydrological behaviour of the sludge (Figure 6.2-9).

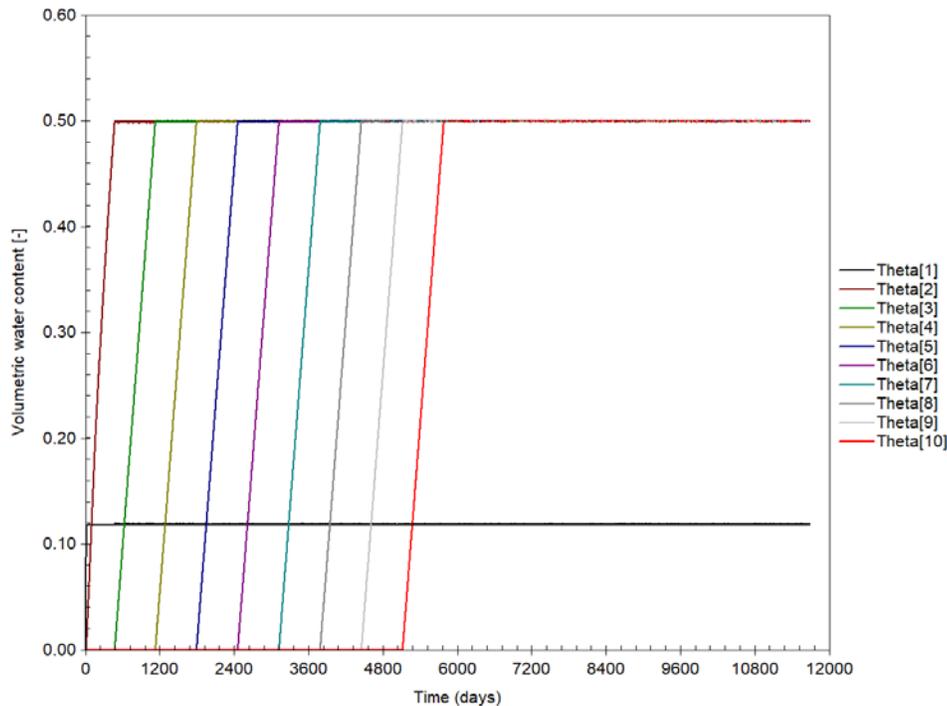


Figure 6.2-9: Soil hydrological profile calculated by ECOFOR in a simulation with average meteorology inputs for the Belgian NORM site. This model run starts with initially dry moisture conditions in the soil. Results expressed as volumetric water content (θ for each soil layer from top (layer 1) to bottom (layer 10)).

The first set of results is the comparison of moisture profiles for the NORM site and the Mol forest, which serves to indicate the differential traits of this unusual sludge soil compared with a more common forest sandy soil. The model shows that in the Mol forest (sandy soil), the high permeability of the sand and its associated high hydraulic conductivity cause a fast response to precipitation events, as water circulates very fast across the heterogeneous 6-layer horizon, draining from the bottom.

The upper soil layer has the highest water content, and is the most time-variable in response to the precipitation changes, and it is also the most saturated. The lower layers have less water content because they have low FC, and θ for each layer is in fact closely around FC due to the fast drainage. By contrast, in the NORM site, the CaF₂ sludge is nearly impermeable and has a low hydraulic conductivity, except for the top soil layer which behaves similarly to the top layer of the Mol forest. The soil layers underneath have a high field capacity and slow drainage, which presents itself in our model simulation as water circulating slowly and θ for each layer being close to FC = 0.5, also close to the porosity (ϵ).

Overall, ECOFOR gives logical predictions with respect to transfer of water to the trees. The water fluxes are higher by an order of magnitude for the NORM site because the water is relatively trapped in the layers whereas in the Mol site the water drains away from the roots more quickly.

We have been able to calculate with ECOFOR realistic soil to plant transfer factors (TFs), defined as the ratio of concentration of the radioelement in whole tree (Bq kg⁻¹) to concentration of element in soil (Bq

kg⁻¹ of soil). To calculate the concentration in whole tree, we summed the total activity roots, wood and leaves, and divided by the tree biomass. To calculate the available radionuclide in Bq kg⁻¹ of soil, we used the sum of activity in Bq for both available and unavailable fractions, divided by mass. Figure 6.2-10 shows TFs calculated by the model in a simulation with average meteorology inputs for the site, including the elements studied in this project (U, Pb and Ra) but also additional elements such as Cl, Ca, K, Mg and Mn and the additional radionuclides Th and Po (calibration for these additional elements and radionuclides performed by a chemical analogue method).

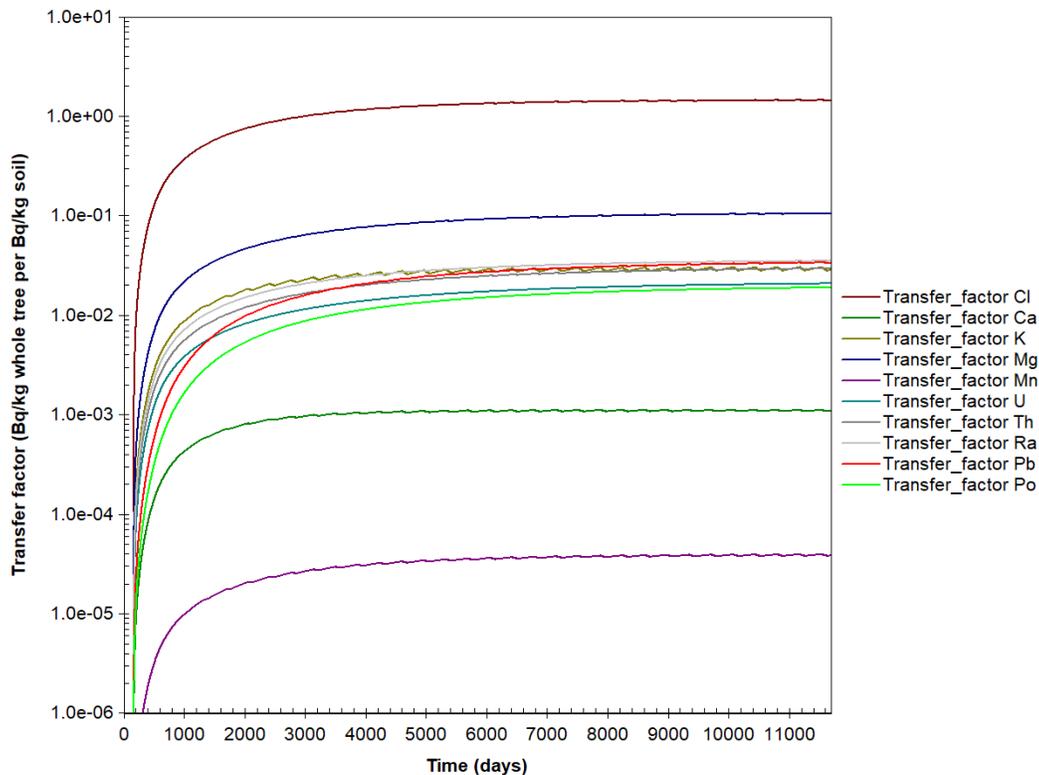


Figure 6.2-10: Soil-plant transfer factors calculated by ECOFOR in a simulation with average meteorology inputs for the Belgian NORM site

The results show that U has a somewhat lower TF than Pb and Ra. We are now in the process of comparing the TF values of our model with these of other models applied in the project TERRITORIES, and for this purpose we have declared a root mean squared logarithmic error (RMSLE) of 1.47×10^{-2} , a bias of 7.05×10^{-4} and a Geometric mean bias (MG) of 1.02 in comparing the TFs of U, Ra and Pb with the experimentally measured values (see below).

6.2.3 TALAL

The simulation results from TALAL reflect the local hydraulic conditions at the NORM site and the hydraulic characteristics of the CaF₂ sludge. The upper part of the sludge (0 – 60 cm) receives water inputs from the atmosphere only (i.e. precipitation) with little influence of the local water table (drainage conditions).

Subsequently, the upper part remains below the field capacity for most of the year (Figure 6.2-11). Changes in moisture are more dynamic in the top 20 cm compared to the deeper layer.

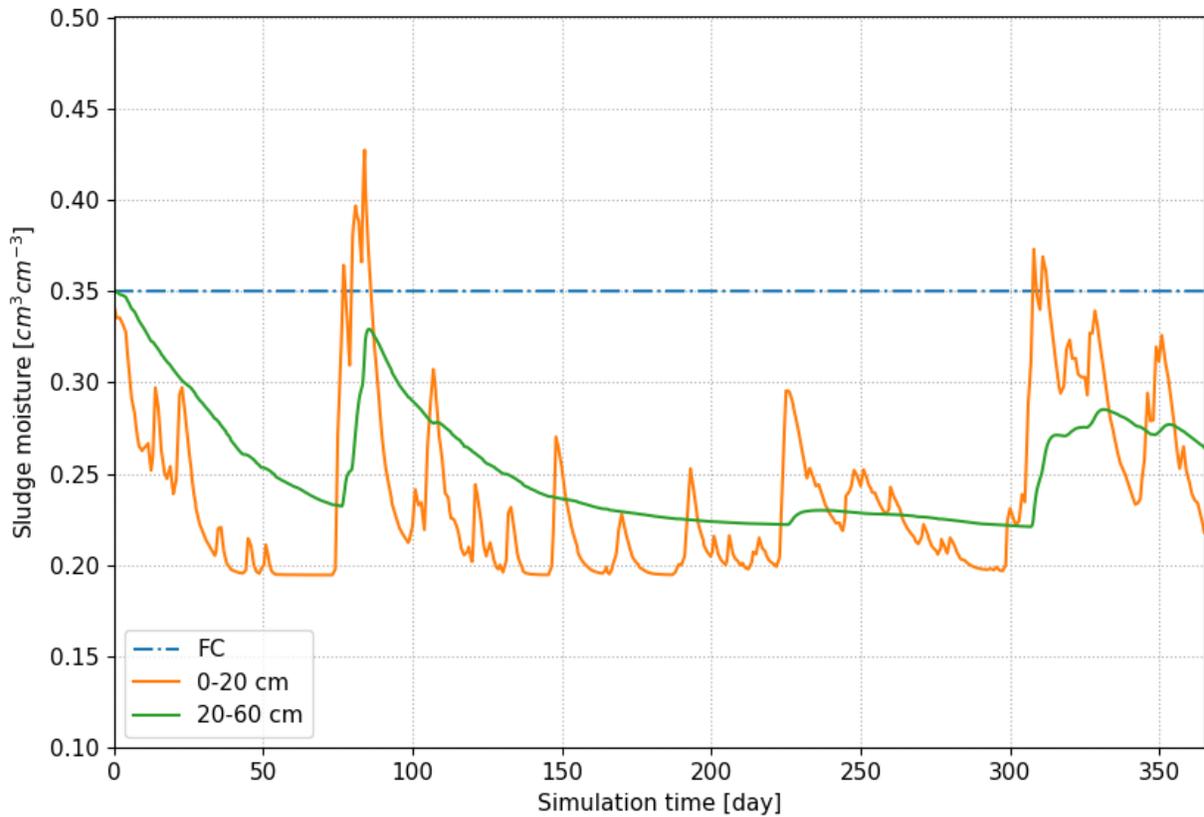


Figure 6.2-11: Simulated dynamics over the course of one year of sludge moisture in response to atmospheric fluxes (rainfall and evaporation) and root water uptake. Simulations reflect a little influence of the local water table, which is located at a greater depth (~ 1 m).

NORM leaching from the sludge is limited due to the slow movement of water (slow advection) and the high sorption (Figure 6.2-12). The decrease in Pb-210 activity concentration in the sludge (Figure 6.2-12) is mainly due to radioactive decay (Pb-210 has a physical half-life of 22 years) rather than leaching.

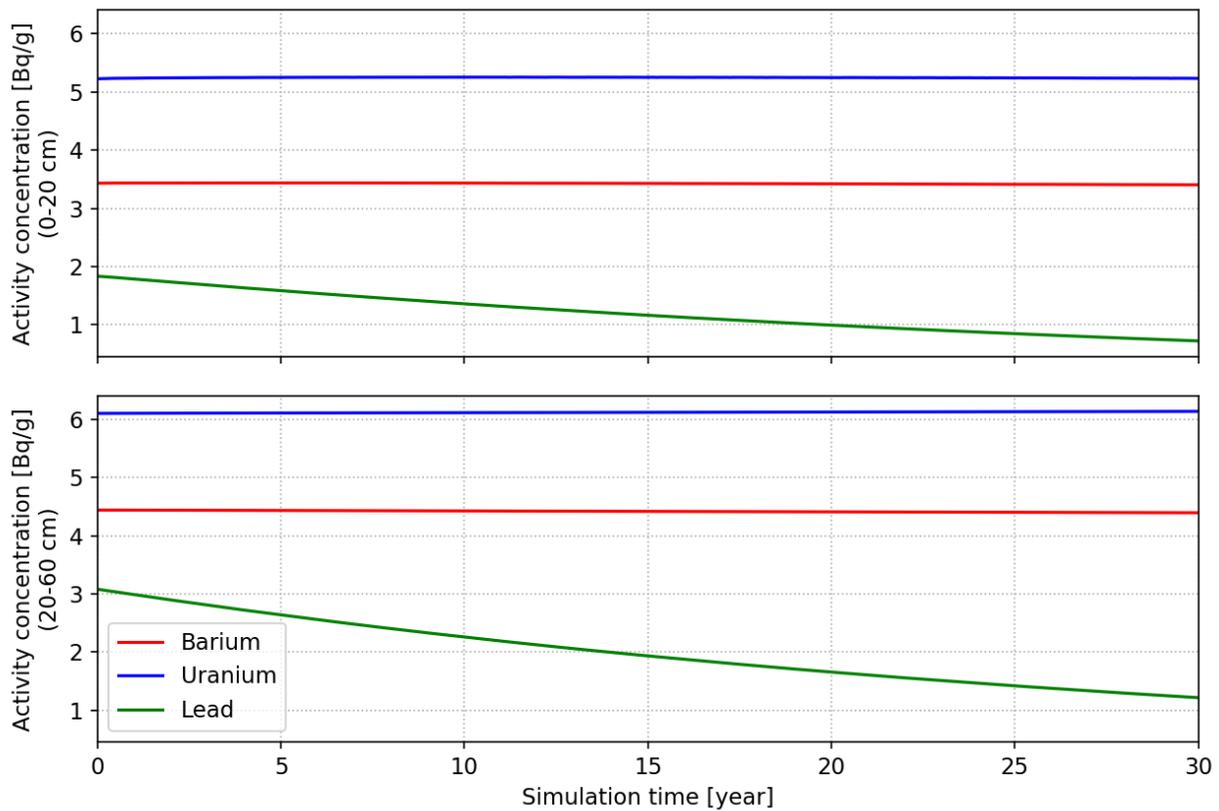


Figure 6.2-12: Simulated NORM dynamics within the sludge.

The NORM dynamics in the tree compartments are shown in. The tree takes up NORM via its roots and accumulates them in its roots, trunk and foliage (Figure 6.2-13). NORM activity concentrations in those compartments reach a quasi-steady state after two decades. Afterwards, lead activity concentration starts to decrease due to radioactive decay and reduced uptake from the root zone.

The transfer factor for NORM followed closely the dynamics of the NORM in the sludge and tree compartments (Figure 6.2-14). The transfer factors increase steadily until they reach a quasi-steady state after two decades. Afterwards, the transfer factor for barium decreases with time.

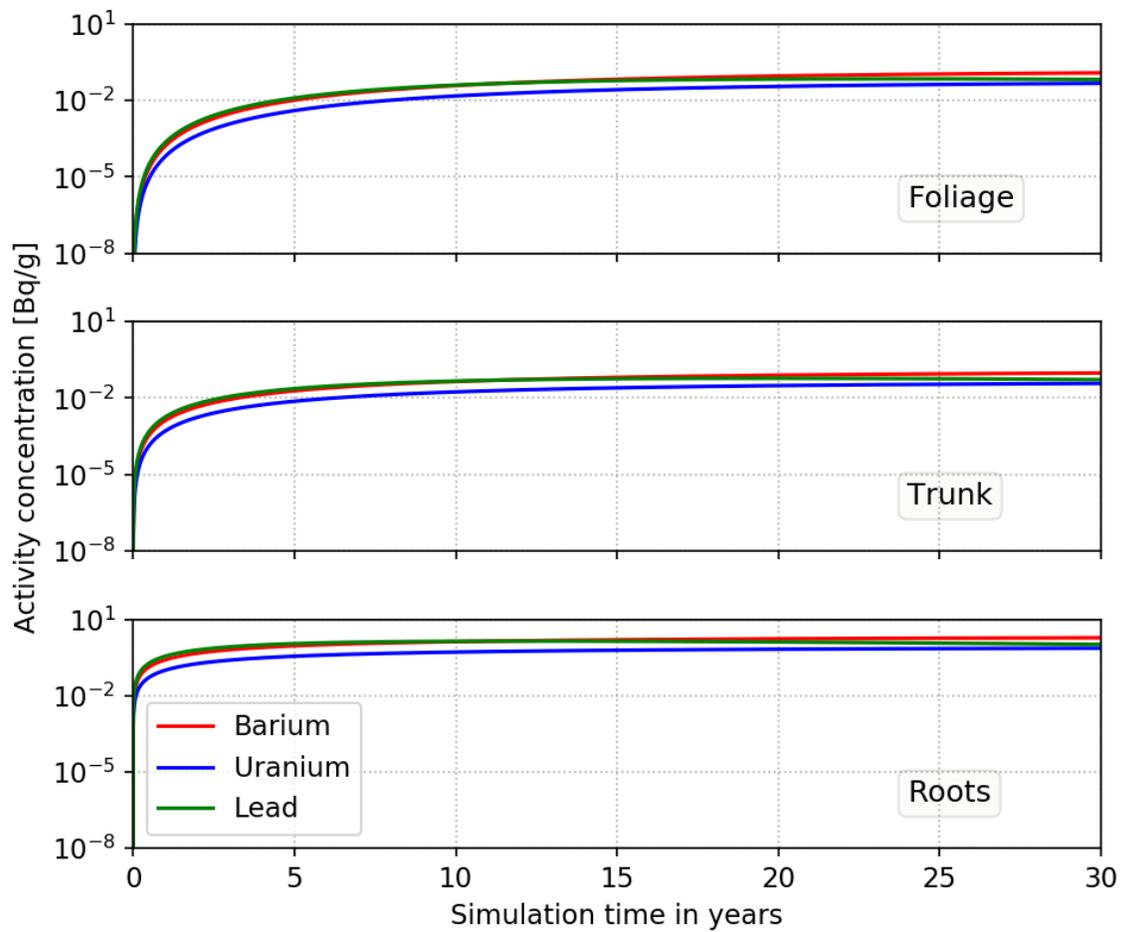


Figure 6.2-13. Simulated dynamics of NORM activity concentration in tree compartments following uptake from the sludge.

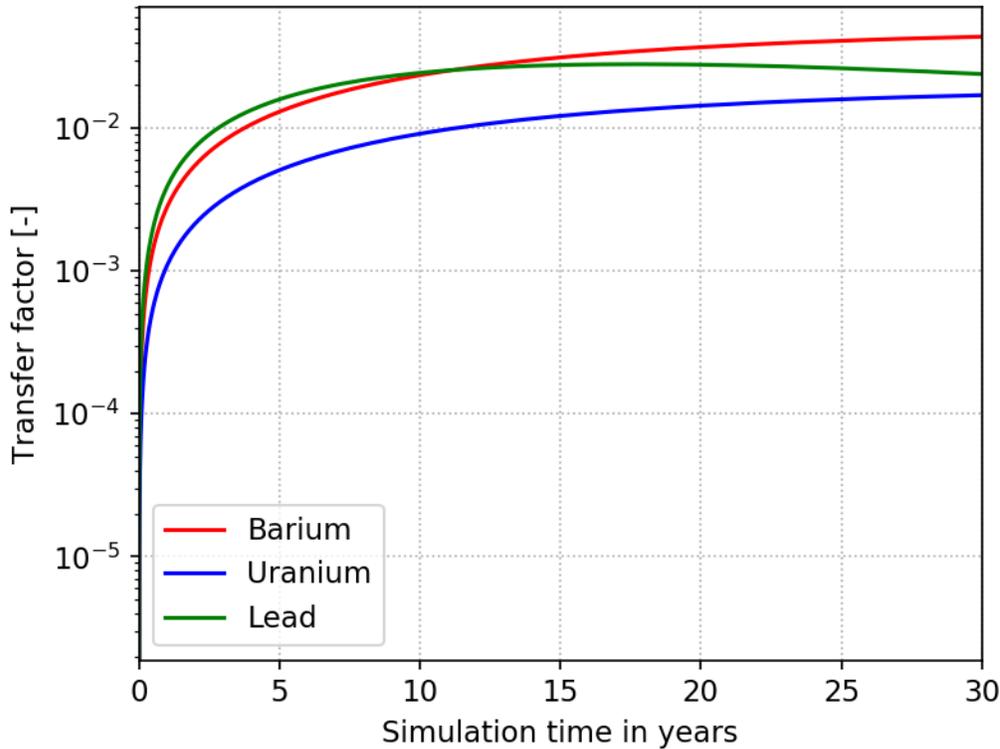


Figure 6.2-14: Simulated transfer factor dynamics for the NORM in CaF₂ sludge.

The simulated transfer factors at the end of the simulation period (30 years) are comparable (same order of magnitude) with the measurements (Table 6.2-6) indicating a good performance of the model. The comparison between measured and predicted transfer factors for Ra should be interpreted with caution since the measured value was obtained by dividing Ra concentration in tree to Ba concentration in sludge.

Table 6.2-6: Measured and predicted (maximum) transfer factors for the NORM.

	Measured	Predicted
U	0.019	0.017
Ra	0.031	0.044
Pb	0.032	0.028

Measured and predicted transfer factor data were compared following the guidance provided in chapter 3.6.3. The root mean squared logarithmic error (RMSLE) was used to provide a measure of model performance:

$$\text{RMSLE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log y_i - \log \hat{y}_i)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(y_i / \hat{y}_i))^2} \quad (6.1)$$

Where:

\hat{y}_i = modelled value; y_i = measured value

A RMSLE of 0.098 was calculated for TALAL based on the 3 pairs of measured-predicted transfer factors in Table 6.2-6.

7 CASE STUDY: FUKUSHIMA

7.1 Introduction to the case study

The terrestrial ecosystems located in the Fukushima region have been exposed to intense radioactive atmospheric fallouts after the FDNPP accident. Due to its long half-life (30 years), ^{137}Cs is of primary importance because it is likely to contaminate these territories for decades, with a long-term impact on humans and biota. Forest ecosystems are particularly vulnerable because these are known to: efficiently capture atmospheric pollution due to their high interception canopy surfaces, recycle pollution along with large amounts of nutrients and store it in the organic soil layers due to their high retention capacities.

To improve our understanding of the biological, physical and chemical processes involved in the cycling of airborne radiocesium deposited onto forests, IRSN conducted a detailed review of literature data acquired from March 2011 to March 2017 at numerous forest sites located in Fukushima and neighboring prefectures. A focus was put on even-aged Japanese cedar (*Cryptomeria japonica*) and Japanese cypress (*Chamaecyparis obtusa*) forests because these were by far the most investigated ones in the contaminated territories. The review covers 82 sites that were classified into 2 categories depending on the dominant tree species which populate the forest stand: evergreen coniferous (EGC, 70 sites) and deciduous broadleaf (DBL, 12 sites). Data on ^{137}Cs activity in tree vegetation and soil layers were carefully compiled, analyzed then processed (see hereafter), before being uploaded in the TLD. The dataset includes time measurements of:

- activity concentrations (in Bq kg^{-1}) in soil layers (e.g. organic and mineral layers every 5 cm depth down to 20 cm) and tree organs (e.g. foliage, branches, trunk bark and trunk wood),
- total activity inventories (in Bq m^{-2}) in soil and tree vegetation,
- tree depuration fluxes ($\text{Bq m}^{-2} \text{d}^{-1}$), with a distinction made between the respective contributions of throughfall, stemflow and litterfall to the total depuration flux.

When stated in the publications, further information regarding the characteristics of the forest sites such as the mean stand age (y), the stand density (trees ha^{-1}) or the above ground biomass (kg m^{-2}) was also considered in the modelling study.

All radioactive quantities were decay-corrected to the sampling date. As radiocesium deposit greatly varies among sites, from less than 10 kBq m^{-2} to more than $1\,000 \text{ kBq m}^{-2}$ to the north-west of the site, radioactive quantities were further normalized by the total deposit (Bq m^{-2}) estimated at each site. This treatment enabled to strongly reduce the spatial variability between sites, although some residual variability persisted which was likely attributed to differences in deposition conditions, differences in forest or climatic characteristics and statistical inaccuracies in the measurements (due to a small-scale heterogeneity under forest canopies). In the TLD, these normalized quantities are referred to as concentrations, fluxes or inventories per unit deposition. Normalized concentrations and fluxes express in $\text{m}^2 \text{kg}^{-1}$ and d^{-1} , respectively, while inventories are unitless numbers (or %). As explained in the CONCERT-TERRITORIES deliverable report D9.59 (Smith *et al.*, 2019), it was further recognized that no single field study was comprehensive enough to provide a complete picture of radiocesium dynamics in Japanese forests over the whole 6-year period. Detailed information on radiocesium concentrations, inventories and fluxes was never available at the same site and over the whole period. This prevented us from carrying a comprehensive modelling study at any site. This is the reason why one adopted a “site-average” approach in which the mean evolution of radiocesium was estimated for each forest category (EGC and DBL) by log-averaging “site-specific” values among the corresponding sites.

7.2 Application of the models to case study

7.2.1 FORESTCROM

The case of Fukushima was assumed to be an instantaneous deposit on trees and soil. This includes an additional uncertainty to the model, as in reality the deposition occurred during several days, but to compare the data obtained for 5 years after the accident, 1 week can be considered as instantaneous. Only Cs-137 was considered, as plenty of information can be obtained from Chernobyl accident and other experiments.

$\lambda_{soil} = \lambda_{phy} + \lambda_{ec-s}$, where $\lambda_{phy} = 6.3258E-5 \text{ d}^{-1}$ is used for Cs-137 (SRS 19) and for Cs-137 a value $\lambda_{ec-s} = 1.4E-4 \text{ d}^{-1}$ was suggested (SRS 19). In the example this $\lambda_{ec-s} = 1.4E-4 \text{ d}^{-1}$, was too small, and it was decided to change to $\lambda_{ec-s} = 1.4E-3 \text{ d}^{-1}$, maximum value provided in the IAEA SRS19 (IAEA, 2001).

For $\lambda_{tree} = \lambda_{phy} + \lambda_{ec-t}$, a $\lambda_{phy} = 6.3258E-5 \text{ d}^{-1}$ can be used for Cs-137 (IAEA, 2001). A λ_{ec-t} can be inferred from TRS-472 data, obtained with Chernobyl data. 3.5 months were derived as average, therefore $\lambda_{ec-t} = 6.51E-3 \text{ d}^{-1}$. For the CR, the CROM (or Erica) provides a CR value for pine trees (conservatively derived) $CR=0.1355$, and therefore $k_{s-t} = 8.9E-4 \text{ d}^{-1}$. For the initial deposition 10.000 Bq/m^2 , the maximum total deposition calculated around Fukushima-Daiichi NPP was assumed (Dvorzhak *et al.*, 2012). A summary of the values considered in this approach is provided in table 7.2-1 and results in Figure 7.2-1.

Table 7.2-1: Values used for a first test of the model.

NAME	VALUE	REFERENCE
$\lambda_{ec-t} \text{ (d}^{-1}\text{)}$	6.5E-03	From IAEA TRS472 (IAEA, 2010) - half lives from 3-4 weeks to 3 months in growing phase, to 4-6 months in dormancy phase.
CR	1.4E-01	From Erica-tool (v. 2015) or CROM (v. 8.3.0)
α	8.0E-01	From IAEA TRS472 (IAEA, 2010) page 100, and Imamura <i>et al.</i> (2017).
deposition (Bq/m ²)	1.0E+05	Dvorzhak <i>et al.</i> (2012)
$k_{t-s}^{out} \text{ (d}^{-1}\text{)}$	4.2E-03	From TRS472 (IAEA, 2010) - equilibrium is reached in 4.5 y (I took $T_{1/2}=1/10$ that value)
$\rho \text{ (kg/m}^3\text{)}$	1.3E+03	Usual soil density from 1300 to 1500 (lower value)
h (m)	2.0E-01	Almost all the activity after 7 years is in the first 10 cm in Fukushima
$\lambda_{soil} \text{ (d}^{-1}\text{)}$	1.4E-04	From SRS19 (IAEA, 2001) pg 65 should be 1.4E-4
$\lambda_{phy} \text{ Cs-137 (d}^{-1}\text{)}$	6.3E-05	Decay constant for Cs-137

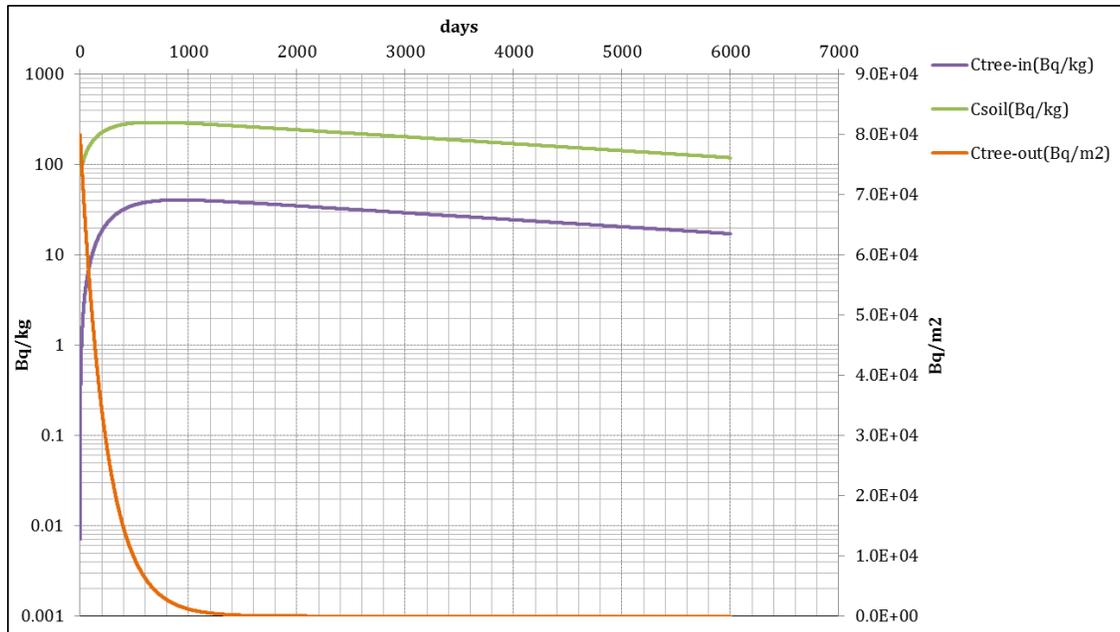


Figure 7.2-1: Results of the model with the parameter values from table 7.2-1.

All aforementioned parameters have a wide range of values which can be applied, and therefore an appropriate uncertainties analysis is needed. Another possibility is given by the values provided in Table 7.2-2 and Figure 7.2-2. In this case the CR is taken as the geometrical mean given in IAEA TRS479 (IAEA, 2014) for trees: coniferous (page 39); the maximum total deposition in the investigated locations is assumed to be $6.3E+05$ (Imamura *et al.*, 2017).

Table 7.2-2: Values used for a second test of the model.

NAME	VALUE	REFERENCE
λ_{ec-t} (d ⁻¹)	6.5E-03	From IAEA TRS472 (IAEA, 2010) - half lives from 3-4 weeks to 3 months in growing phase, to 4-6 months in dormancy phase.
CR-GM	7.5E-02	From IAEA TRS479 (IAEA, 2014) table 5, page 39. Trees coniferous.
CR-GSD	3.2E+00	From IAEA TRS479 (IAEA, 2014) table 5, page 39. Trees coniferous.
CR-Min	1.2E-03	From IAEA TRS479 (IAEA, 2014) table 5, page 39. Trees coniferous.
CR-Max	1.8E+00	From IAEA TRS479 (IAEA, 2014) table 5, page 39. Trees coniferous.
α	8.0E-01	From IAEA TRS472 (IAEA, 2010) page 100, and Imamura <i>et al.</i> (2017).
deposition (Bq/m ²)	6.3E+05	Imamura <i>et al.</i> (2017).
k_{t-s}^{out} (d ⁻¹)	4.2E-03	From TRS472 (IAEA, 2010) - equilibrium is reached in 4.5 y (I took T1/2=1/10 that value)
ρ (kg/m ³)	1.3E+03	Usual soil density
h (m)	2.0E-01	Almost all the activity after 7 years is in the first 10 cm in Fukushima
λ_{soil} (d ⁻¹)	1.4E-04	From IAEA SRS19 (IAEA, 2001) pg 65 should be 1.4E-4.
λ_{phy} Cs-137 (d ⁻¹)	6.3E-05	Decay constant for Cs-137

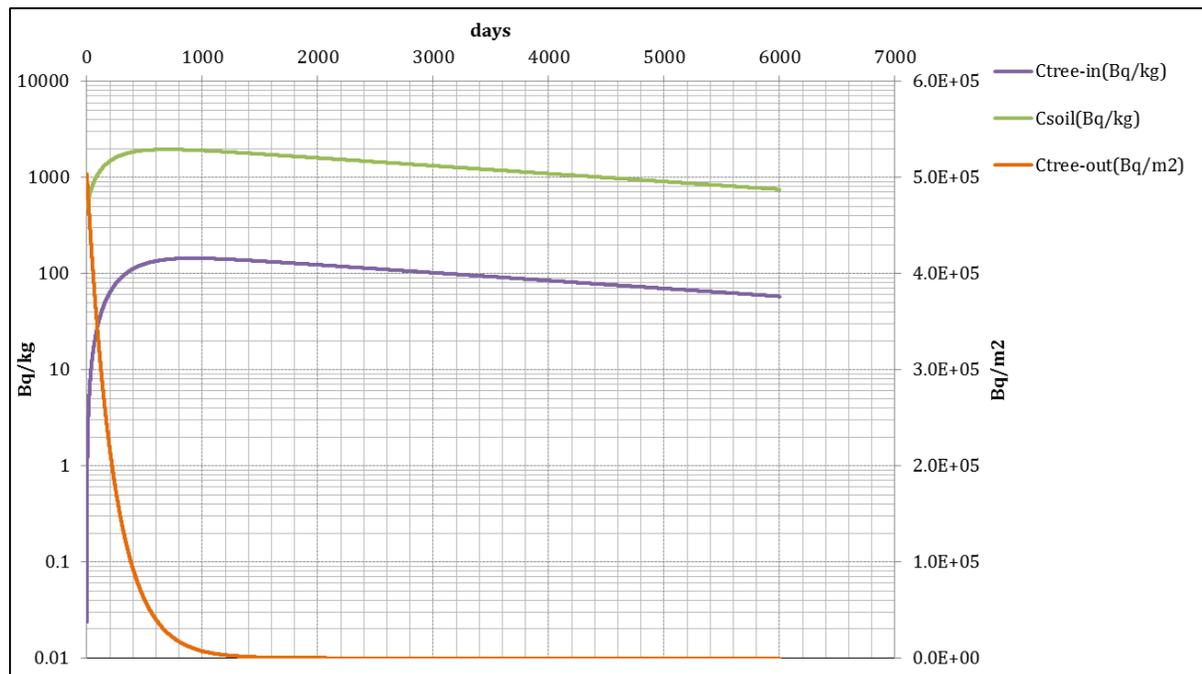


Figure 7.2-2: Results of the model using the parameters from table 7.2-2.

Many other combinations of the values for the different parameters can be tested from previous studies, basically based in Chernobyl studies (for instance, the interception factor can be selected from 0.4 to 0.6 for pine forests 30 years old, to 0.9 to 1 for 6 to 10 years old pine forests), average soil density strongly depends on the organic matter in the soil and therefore the age of the forest, the different ecological parameters must be adapted to the climatic conditions of the location, etc.. This deserves further research for the selection of appropriate parameters in specific conditions of Fukushima pine forests and an uncertainty and sensitivity analysis to obtain the range of values obtained as a result of the concentrations in the soil and the inner part of the trees, which are well documented for different times after the Fukushima accident.

In any case some conclusions can be extracted from the tests. As can be seen in the two simulations, equilibrium between soil and trees (and therefore the application of a simple CR approach is valid) is achieved in the pine trees after around 600 to 800 days (2 to 3 years) after the accident. The contribution of the external contamination of the tree to the contamination of the soil can be considered negligible after nearly 4 years due to the calibration of the model using Chernobyl data. The difference between the use of a constant concentration in the soil, calculated using the maximum total deposition and no interception factor (2420 Bq/kg), and a conservative CR (0.14) which gives a maximum concentration in the inner part of the tree of 340 Bq/kg, to a more refined approach using the described model and more precisely selected values for the different parameters (maximum 155 Bq/kg (Imamura, *et al.*, 2017)) results in a conservatism of a factor 2.

More site specific information can be extracted from recent articles published with measurements in Fukushima. According with Imamura *et al.* (2017), the total deposition for Cs137 in one of the investigated sites (KU1-S was used because the maximum deposit was in this point) was 630 kBq/m². Moreover, the needles ecological half-life was provided for each site, being 0.7 y for KU1-S (which corresponds to a k_s^{out} of $2.7E-03 d^{-1}$). Additionally the biomass for needles was determined (from 1.2 to 1.4 kg/m² for needles in conifers at KU1-S). If the model is calibrated using the input value (based in aerial survey measurements), and those 2 locally determined parameters, leaving all the other as used above (based in studies basically from Chernobyl), as shown in table 7.2-3, the results obtained in the figure 7.2-3 are obtained.

Table 7.2-3: Values used for the parameters of model (some of them locally determined).

NAME	VALUE	REFERENCE
λ_{ec-t} (d ⁻¹)	6.5E-03	From IAEA TRS472 (IAEA, 2010) – half-lives from 3-4 weeks to 3 months in growing phase, to 4-6 months in dormancy phase.
CR-GM	1.4E-01	From Erica-tool (v. 2015) or CROM (v. 8.3.0)
CR-GSD	3.2E+00	From IAEA TRS479 (IAEA, 2014) table 5, page 39. Trees coniferous.
CR-Min	1.2E-03	From IAEA TRS479 (IAEA, 2014) table 5, page 39. Trees coniferous.
CR-Max	1.8E+00	From IAEA TRS479 (IAEA, 2014) table 5, page 39. Trees coniferous.
α	0.457	Calibrated using data from Imamura <i>et al.</i> (2017)
deposition (Bq/m ²)	6.3E+05	From Imamura <i>et al.</i> (2017). (KU1-S)
k_{t-s}^{out} (d ⁻¹)	5.4E-04	From Imamura <i>et al.</i> (2017). (KU1-S)
ρ (kg/m ³)	1.3E+03	Usual soil density
h (m)	2.0E-01	Almost all the activity after 7 years is in the first 10 cm in Fukushima
λ_{soil} (d ⁻¹)	1.4E-04	From IAEA SRS19 (IAEA, 2001) pg 65.
λ_{phy} Cs-137 (d ⁻¹)	6.3E-05	Decay constant for Cs-137
Biomass needles (kg/m ²)	6.3E-05	Decay constant for Cs-137

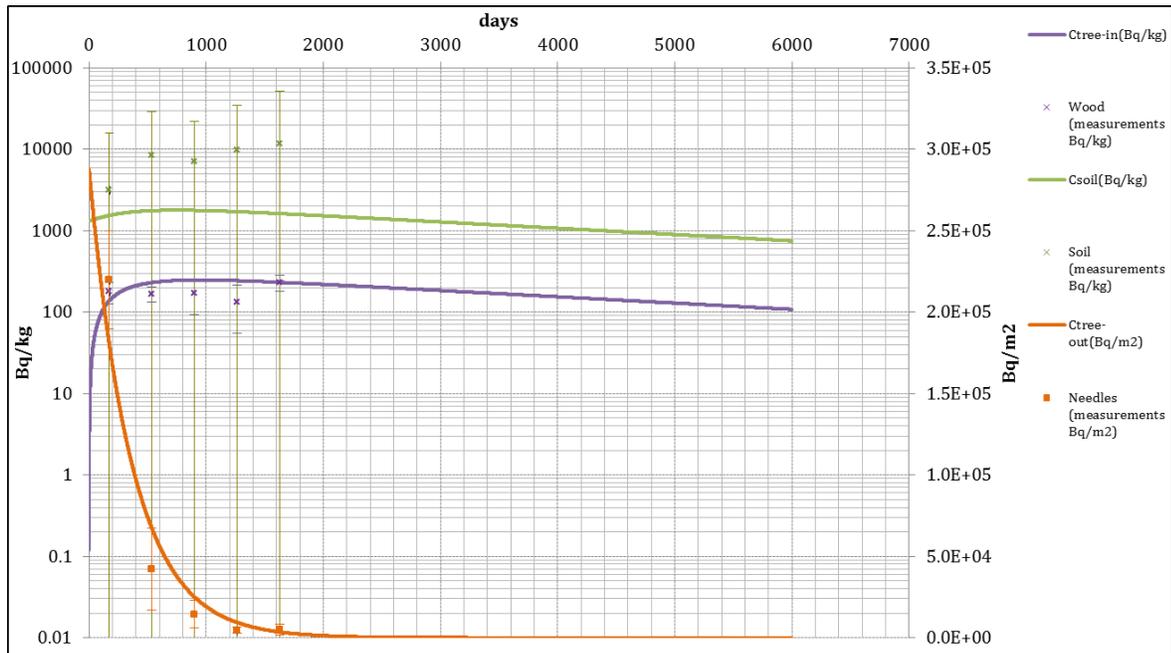


Figure 7.2-3: Results of the model using the parameters from table 7.2-3 (log scale).

In this case, as can be seen, the change of a single locally determined factor as is the case of the washout of the external contamination of the trees; imply drastic changes in the results of the model. The more important change occurs in the time where the maximum concentrations in soil and trees are predicted 7-8 years after the accident, achieving the tree (as a whole organism) almost a constant activity concentration for the 20 years after the accident. Measured data of wood concentrations in the trees as a whole in Fukushima (KU1-S) (Imamura *et al.*, 2017) were in the range of 134 – 230 Bq/kg for the period August 2011 to August 2015, being the values almost constant. The maximum concentration predicted with the model was 242 Bq/kg, what is somehow overestimating the real values.

RMSLE obtained in this case was 0.22 for needles, 0.16 for wood and 0.67 for soil. As aggregated value, RMSLE was calculated to be 0.72.

Some more refinements in the parameters used in the model, including a temporal dependence of the average depth of the concentration in soil, and a temporal dependence of the CR with the age of the tree, would probably improve even more the results of this extremely simple model.

7.2.2 TREE4-simple

In this section, we summarise the results obtained with TREE4-simple model for ^{137}Cs in Fukushima forests. The performance of the model was evaluated by comparing predicted values with time observations over the period 2011-2017 for some selected endpoints. The forests sites under consideration are even-aged Japanese cedar and cypress stands (i.e. coniferous forests), for which the dataset is rich and reliable.

Estimation of model parameters

Regarding the estimation of model parameters, most of the “easily measurable” parameters could be estimated based on a literature review of field studies conducted before or after the Fukushima accident. This analysis enabled to estimate, with more or less confidence, a probability density function for each of them, as well as correlations in-between some of them. As a general rule, we imposed (log-) uniform pdfs for parameters with little knowledge and (log-)normal pdfs when more information was available. The data collected from literature clearly demonstrated that there were no significant differences in the eco-physiological or hydrological characteristics between the two tree species. We thus adopted the same parameter values for the two species. Uncertainties in the annual climatic characteristics were estimated by analysing meteorological data collected within a few hundreds of kilometres from the FDNPP.

The few remaining unknown (or “hardly measurable”) parameters of the Cs module, and correlations in-between, could not be estimated otherwise than by calibrating the Cs module outputs (i.e. concentrations, inventories and fluxes) against the 6-year observations. The calibration step concerns the following parameters: the throughfall rate, the incorporation rate, the effective Kd of the litter layer and the fixation rate in the soil layer. Calibration was carried out by minimizing the mean squared errors specified hereafter.

Methodology

We performed Monte-Carlo simulations (10^4 runs, LHS sampling) taking into account parameter uncertainties, and correlations in-between them based on the Iman and Conover method. The simulations were run for the period from March 2011 to March 2017, with an initial ^{137}Cs deposit of 1 Bq m^{-2} . To quantitatively assess model performance, the difference between (probabilistic) model outputs and measured value was estimated through the calculation of the (probability density function of) the Mean

Squared Error (MSE), or Logarithmic Mean Square Error (MSLE), integrated over the observation period. These metrics were computed for the variety of forest components and/or mathematical outputs which could be compared with observations.

For the total activity inventories (unitless) in soil and tree vegetation, which vary from 0 to 1, we computed an aggregated MSE combining both outputs as follows:

$$MSE = \sum_i \alpha_i \frac{1}{T_i} \int_{T_i} (y_i^{mod} - y_i^{obs})^2 dt$$

↑ Cs inventory in i
↑ observation periods for i
↑ weighting factor for i (=0.5)
↑ i=tree, soil

(7.1)

Where $y_i^{mod}(t)$ and $y_i^{obs}(t)$ are the predicted and measured activity inventories in i =soil or tree, and α_i are weighting factors fixed at 0.5 in this study. The integration is made over the respective observation periods T_i , which can differ between the 2 compartments. This index quantifies the capability of the model to accurately predict the redistribution over time of the initial ^{137}Cs deposit between soil and tree, i.e. the overall balance between tree depuration and root-uptake.

For the activity fluxes (in d^{-1}) which were measured at field – throughfall, stemflow, litterfall and the sum of them - we computed an aggregated MSLE because two of them decrease by more than 2 orders of magnitude over the observation periods. Using the MSE is not recommended here because we do not want to put more emphasis on the short term than the long term in the quality assessment. The aggregated metric was calculated as follows:

$$MSLE = \sum_i \alpha_i \frac{1}{T_i} \int_{T_i} \left(\ln \frac{y_i^{mod}}{y_i^{obs}} \right)^2 dt$$

↑ Cs flux i
↑ i=throughfall, stemflow, litterfall, total

(7.2)

Where the 4 weighting factors were fixed at 0.25, so as not to give preference to any particular flux. With this index, we estimate the model accuracy in predicting the total depuration flux and the individual contributions of the 3 underlying transfer mechanisms.

For the same reasons as those explained above, we used the MSLE for activity concentrations, also. For tree vegetation, we computed an aggregated index combining errors in foliage, branches, trunk wood and trunk bark, with all organs given equal weight (0.25). For soil, we computed an aggregated index combining errors in the litter and the 0-20cm mineral layer, with equal weight (0.5).

Scenario description

The simulations were carried out for a variety of scenario, depending on the type of plantation (i.e. stand age at time of deposition) and the type of atmospheric deposits (i.e. dry/wet deposition ratio, rainfall height during deposition). In this document, we will focus on a scenario which is quite representative of the contaminated forest sites investigated after the accident. In this scenario, we consider that ^{137}Cs deposition occurred on March 15 (2011) at a constant deposition rate (1 Bq d^{-1}) under low precipitation (2 mm height). In fact, the model results were not significantly affected by the duration of the deposition event, which likely varied between a few hours and a few days at Fukushima sites. The results were shown

to be much more sensitive to the precipitation height, as long as precipitation exceeded 3 to 4 mm. The stand age is in the range 40 ± 10 years (normal law), which is quite representative of the forest sites investigated in Fukushima. As ageing is neglected in TREE4-simple, the age remains constant over the 6-year simulation period, as well as other stand characteristics which are derived from such as biomasses and leaf area indexes.

Results

The probabilistic model outputs are compared with field observations in Figure 7.2- and Figure 7.2-5.

The levels of contamination predicted in the soil-tree system, and their time evolution, are globally consistent with observations, apart from: (i) the activity concentration in the canopy which is strongly overestimated after 2012 (with respect to measured concentrations in needles and branches) and (ii) the activity concentration in trunk wood which is under-predicted before 2012-2013. Agreement on the remaining radiological quantities was met in part thanks to the use of calibrated values for the unknown (sensitive) parameters, i.e. throughfall and incorporation rates which partly determine the tree-to-soil transfer of ^{137}Cs due to depuration during the first 2 years after the accident and litter K_d which strongly influences the rate of migration of ^{137}Cs down to the soil layer. These results suggest that the modelling approach is likely complex enough for satisfactorily predicting the redistribution of the initial ^{137}Cs deposit between the soil and tree compartments, as well as its redistribution between the litter and mineral layers. On the contrary, TREE4-simple seems inappropriate for (further) predicting the activity partitioning between the various tree organs whatever the model parameters are. The model is also able to predict the vertical activity profile in the soil layer as it relies on a single layer assumption.

The data and the underlying mechanisms explaining the dynamics observed in Figure 7.2-4 and Figure 7.2-5 will be further discussed in the Section 7.2.3 dealing with the application of TREE4-advanced.

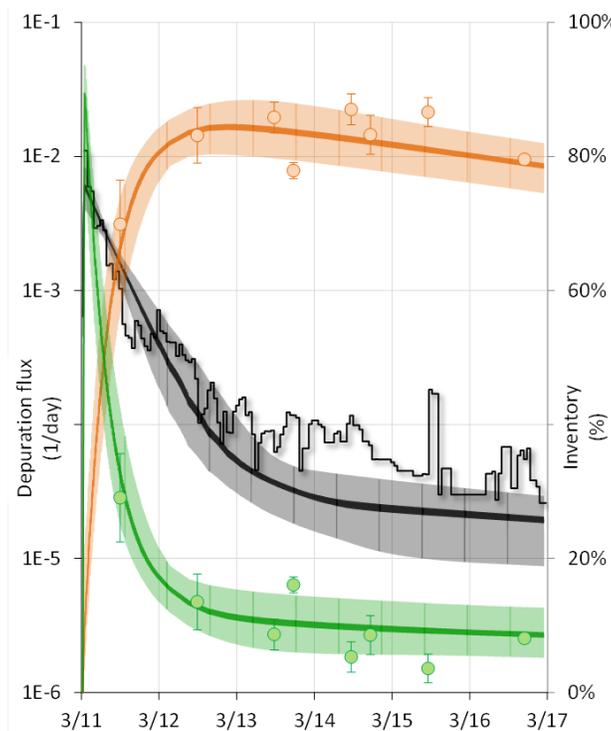


Figure 7.2-4: Predictions versus observations of ^{137}Cs activity inventory in soil, activity inventory in tree and activity depuration flux. The coloured ribbons correspond to the 5%-95% confidence interval.

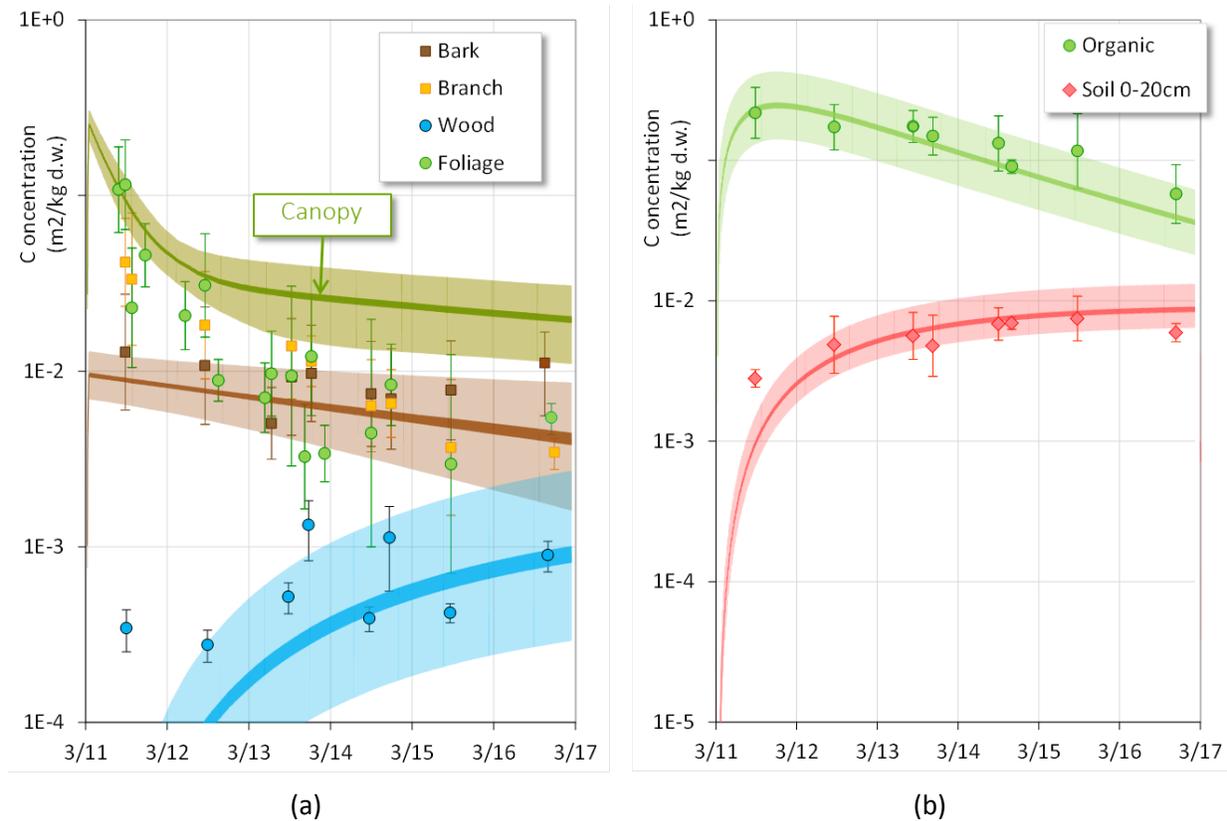


Figure 7.2-5: Predictions versus observations of: (a) ^{137}Cs activity concentrations in the canopy, trunk bark and trunk wood, (b) ^{137}Cs activity concentrations in the organic (i.e. litter) and 0-20 cm mineral soil layers. The coloured ribbons correspond to the 5%-95% confidence interval.

7.2.3 TREE4-advanced

In this section, we summarise the results obtained with TREE4-advanced model in Fukushima forests, adopting the same methodology and scenario as described above for TREE-simple.

Estimation of model parameters

The “easily measurable” parameters newly introduced in TREE4-advanced were estimated based on a literature review, while the unknown parameters, including those already introduced in TREE4-simple, were (re-)estimated by calibrating the eco-physiological and Cs module against observations. The hydrological module did not require any calibration.

The parameters involved in the eco-physiological module were estimated by calibrating time outputs against hundreds of data collected from forestry literature for stand ages ranging from 10 to 70 years (see figure 7.2-6). These kinds of growth curves allow us to predict the ageing of the forest stand characteristics. This is particularly important for long-term simulations (i.e. over decades) because the relative changes in the stand density or wood biomass become significant. In the present scenario, where the stand age increases from 40 ± 10 to 46 ± 10 years, this is less sensitive. In six years, the eco-physiological module predicts that: the stand density decreases by less than 20%, the tree dimensions only slightly increase by about 10% while trunk and branch biomasses increase by about 15%. The influence of ageing would have been greater for a 10-yr old forest stand whose characteristics evolve faster than those of a mature forest stand.

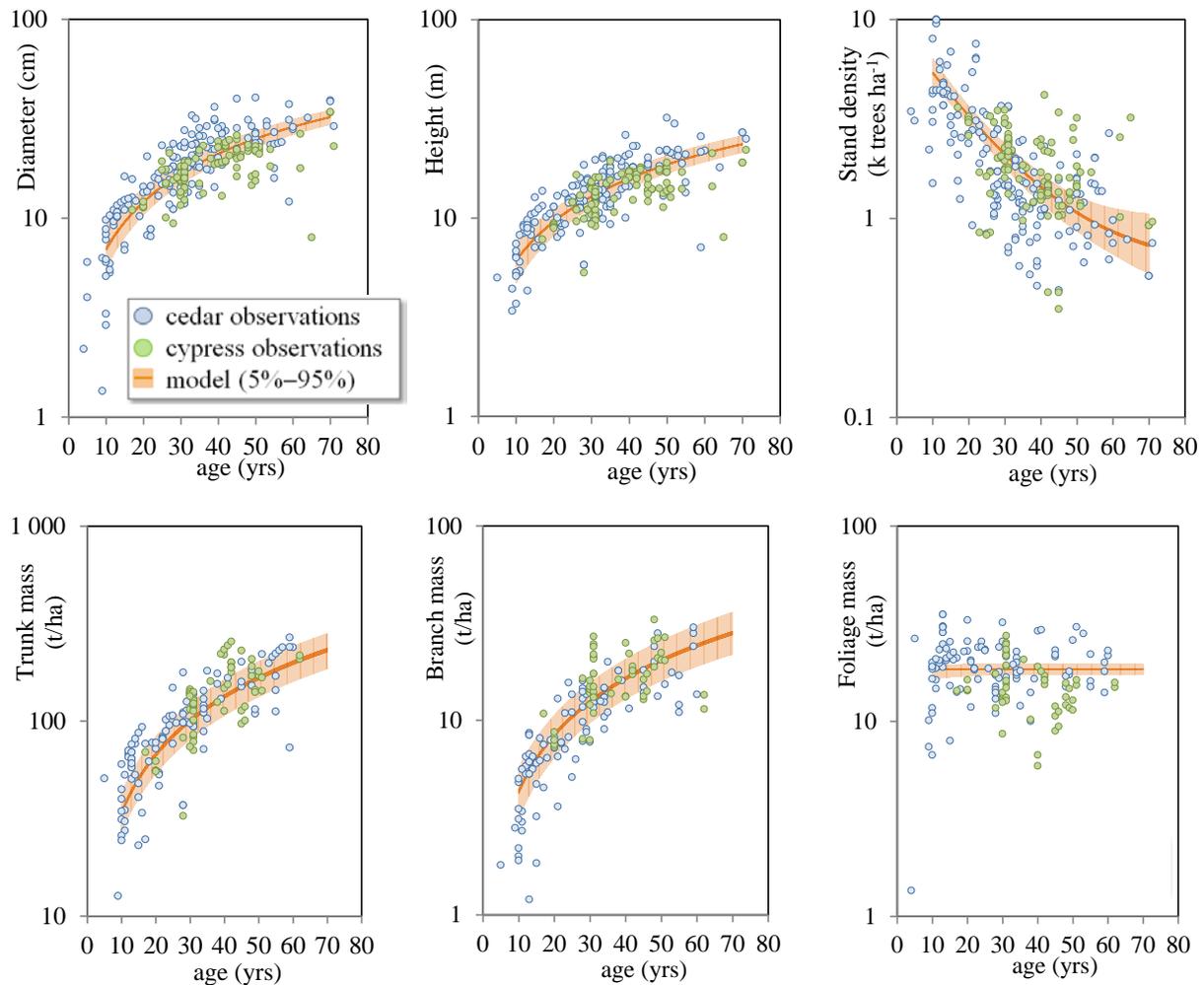


Figure 7.2-6: Predicted versus observed evolution of some eco-physiological characteristics of even-aged Japanese coniferous plantations (cypresses and cedars).

Like for TREE4-simple, the “unknown” Cs module’s parameters, as well as correlations in-between, were estimated by calibrating the module outputs against the 6-year observations based on the minimization of mean squared errors. The calibration step concerns the following additional parameters: the foliage-to-wood translocation rate and the effective Kd of the mineral layer. The two following long-term processes were neglected because they did not play a role over the 6-year simulation period: wood-to-foilage translocation and remobilization.

Methodology

The performance of TREE4-advanced was evaluated using the same metrics (MSE, MSLE) as those introduced for TREE4-simple, based on probabilistic calculations.

Scenario description

As mentioned above, the scenario is identical to that studied with TREE4-simple.

Results

The probabilistic model outputs are compared with field observations in Figure 7.2-7 and Figure 7.2-8.

The levels of contamination predicted in the soil-tree system, and their time evolution, are globally consistent with observations, although predicted variability is strongly underestimated. The accuracy of the predicted activity concentrations in tree organs has been significantly improved compared with TREE4-simple. These encouraging results suggest that the modelling approach is likely complex enough for accurately reproducing the observed dynamics, as long as relevant parameter values are used. It also suggests that such an advanced model might help us get further insights into the understanding of the system and quantify the respective contributions of the various dynamic processes.

As shown in Figure 7.2-7a, the decrease of ^{137}Cs inventory in standing vegetation starts immediately after the accident and obeys two characteristic half-lives, equal to approximately 50 days (during the first 6 months) and 20 months (after 2 years). In the initial stage, tree depuration is mostly driven by throughfall, i.e. weathering of the ^{137}Cs pool initially intercepted by the canopy but not yet incorporated into internal tissues. On the longer term, tree depuration is sustained by litterfall (of dead needles, mainly) and the weathering of ^{137}Cs excreted by foliage. From Figure 7.2-7b, we can notice that litterfall flux is significantly under-estimated by the model; this is likely due to the underestimation of the biomass turnover rate. Although they are far less important than depuration, the model suggests that tree mortality and root uptake start to play a role after 3 to 5 years, typically. The strong infra-annual variability observed on the depuration fluxes could not be reproduced by the model because TREE4-advanced does not account for seasonal or even shorter effects. More generally, it can be seen that predicted variability/uncertainty is too low for all radiological quantities. The reasons are that we did not account for: (i) infra-annual effects, (ii) scenario uncertainties (as the scenario was fixed), (iii) structural uncertainties (as the model was fixed), (iv) uncertainty on the estimated ^{137}Cs deposit (as it was fixed to 1 Bq m^{-2} in our calculations).

As shown in Figure 7.2-8a, predicted activity concentrations in March 2011 in tree organs that are directly impacted by the atmospheric fallouts vary widely, i.e. foliage > branches > trunk bark. These initial concentrations are mainly determined by the area indexes and the specific area coefficients because these two parameters determine the proportion of the deposit which is intercepted per unit of mass of each organ. Concentrations in woody organs decay exponentially with time, with a characteristic half-life equal to ~ 18 months in branches and ~ 10 years in trunk bark. The decrease of ^{137}Cs concentration in foliage obeys to two characteristic half-lives: ~ 2 months (due to throughfall) and 2 years (due to biomass turnover and foliar excretion). Let us remember that tree mortality does not have any incidence on concentrations. Activity concentration in wood quite rapidly reaches an “asymptotic” value equal to $\sim 0.001 \text{ m}^2 \text{ kg}^{-1}$ dry mass. The simulations indicate that the rapid increase in wood is essentially driven by the translocation of ^{137}Cs through the phloem, once incorporated into the internal foliar tissues. The contribution of root uptake in wood contamination at the end of the simulation period was estimated to 30%, not more.

The rate of decrease of activity concentration in the top organic layer (see Figure 7.2-8b) is governed by a competition between litterfall (input) and leaching (output); the decrease seems a little bit over-estimated by the model for reasons that have not been identified yet. The rate of transfer of labile ^{137}Cs from the organic layer to the underlying top 5cm mineral horizon is governed by the percolation flux (estimated to $\sim 1000 \text{ L m}^{-2} \text{ y}^{-1}$) and the organic layer's K_d , the value of which was estimated to $\sim 800 \pm 200 \text{ litter/kg}$ dry mass. The time evolution of the vertical distribution of ^{137}Cs in the mineral profile is difficult to predict because it depends on various physico-chemical and biological processes which are not accounted for in TREE-advanced.

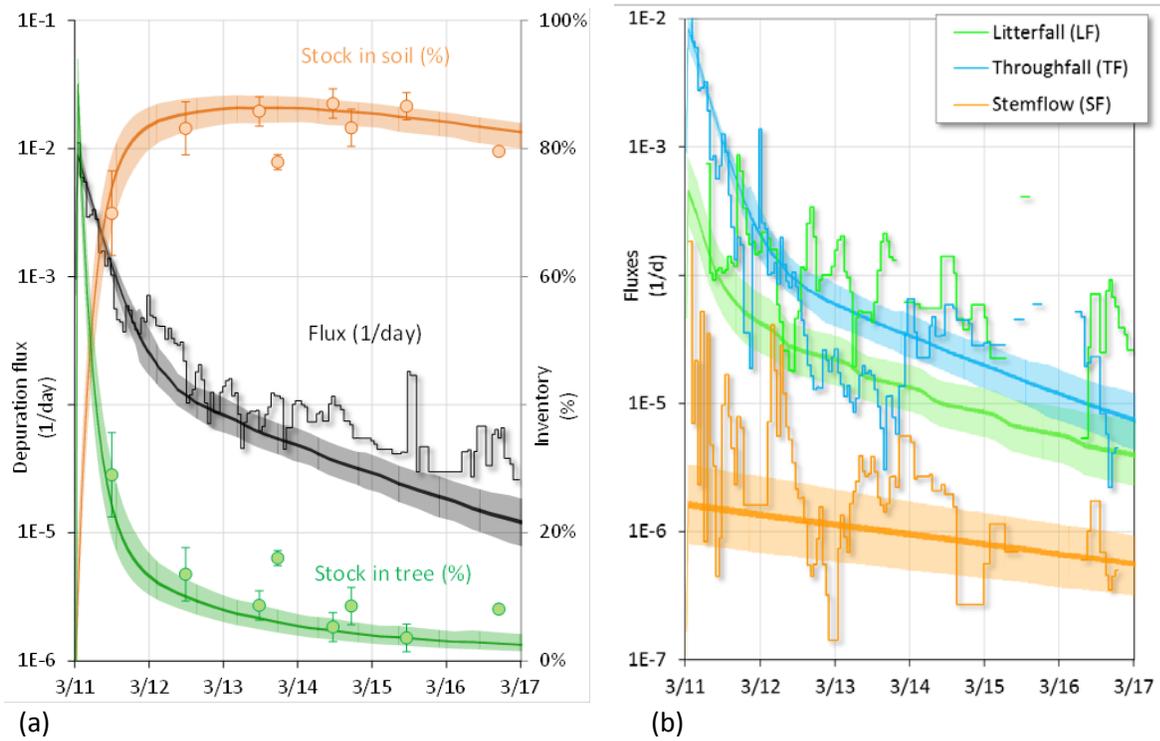


Figure 7.2-7: Predictions versus observations of ^{137}Cs : (a) activity inventory in soil, activity inventory in tree and activity depuration flux, and (b) litterfall, throughfall and stemflow fluxes. The coloured ribbons correspond to the 5%-95% confidence interval.

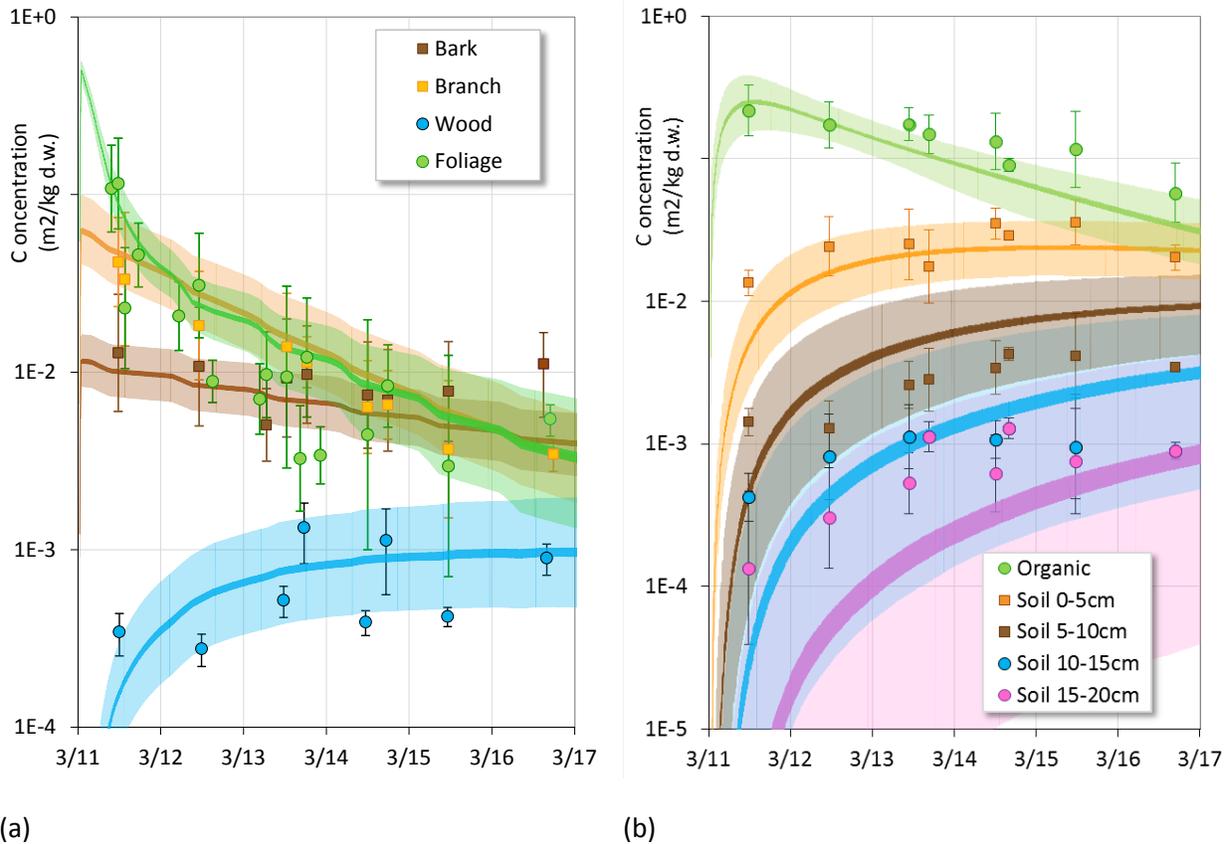


Figure 7.2-8: Predictions versus observations of ^{137}Cs : (a) activity concentrations in the canopy, trunk bark and trunk wood, (b) activity concentrations in the organic layer (i.e. litter) and mineral soil layers every 5 cm depth. The coloured ribbons correspond to the 5%-95% confidence interval.

8 CASE STUDY: SELLAFIELD

8.1 Introduction to the case study

Since the early 1950s radionuclides have been released into the Irish Sea region from routine operations at the nuclear reprocessing facility of Sellafield site, UK. The bulk of this material was released in solution leading to dispersion in the marine environment both in the vicinity of the release and further afield. However, some radionuclides became incorporated onto particles, both natural and manmade, and were subsequently dispersed in particulate form.

Monitoring of ambient levels of radiation has been carried out at the site for many years and measurements are available of radioactivity concentrations in water, sediments and biota. For the TERRITORIES project such data have been compiled for the radionuclides ^3H , ^{99}Tc , ^{137}Cs , $^{239/240}\text{Pu}$ and ^{241}Am and used for model validation.

In addition, large area beach monitoring has taken place near Sellafield since 2006 and has led to the detection of radioactive particles in sand and sediments. The particulate finds were classified in terms of physical size ('objects' > 2 mm, 'particles' < 2 mm) and radioactivity ('alpha-rich' ^{241}Am activity > ^{137}Cs

activity or 'beta-rich' ^{137}Cs activity $> ^{241}\text{Am}$ activity). The location and date of these finds were carefully recorded so trends in their distribution along the Sellafield beaches are known.

This dataset was also made available to the TERRITORIES project and was used in the validation of a model that predicts the dispersion of particulates in the marine coastal environment.

The beach areas near the Sellafield nuclear reprocessing facility, UK, have been exposed to radionuclides released over many decades. Here the possibility to use a simple modelling approach for prediction of the consequences to the beach region during long time of exposure will be evaluated.

8.2 Application of the models to case study

8.2.1 ARCTICMAR and AMIS models. Quantifying improvement the local model (AMIS) versus the NRPA box model (ARTICMAR)

8.2.1.1 Results of the ARTICMAR model compared with measured data

Figures 8.2-1 and 8.2-2 shows results of the ARTICMAR model for water (the initial compartment for the releases from the Sellafield facilities). The release corresponds to liquid discharges of ^{137}Cs from 1952 to 2016 years.

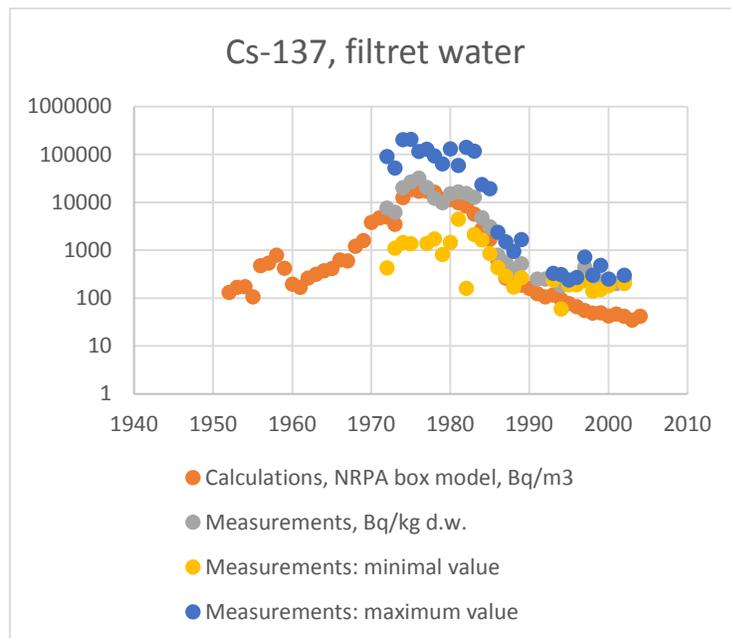


Figure 8.2-1: Filtered waters in the compartment 204 CW1 – Cumbrian waters.

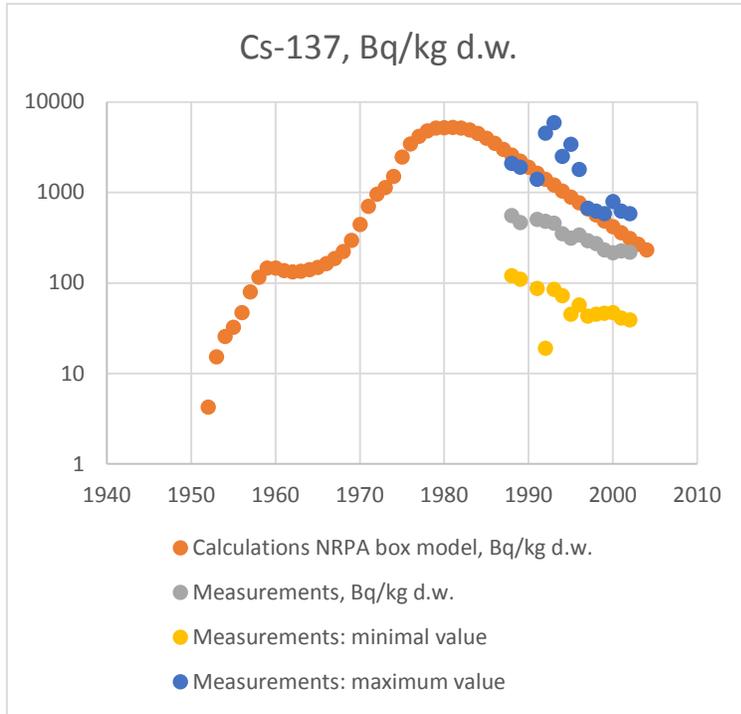


Figure 8.2-2: Sediment in the compartment 205 CW1_sediment –Cumbrian waters.

Figure 8.2-3 shows simulations of the ARTICMAR model for the fallout after Chernobyl accident of Cs-137 for the Finland Gulf (the Baltic Sea).

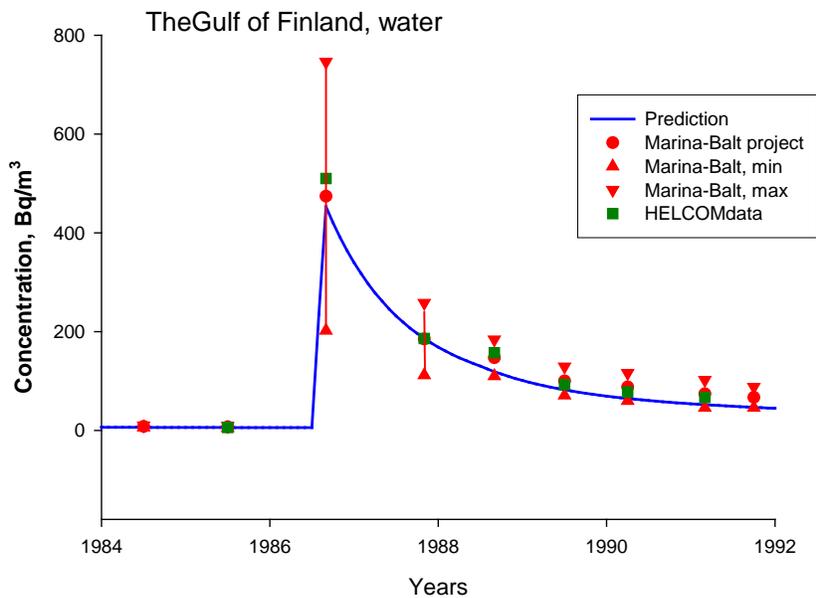


Figure 8.2-3: Filtered waters in the compartment 323 GF1 – Gulf of Finland.

Figure 8.2-4 shows simulations of the ARTICMAR model for global fallout of Pu-239+240 for some coastal regions near Iceland. Figure also shows a comparison of the model simulations with monitoring data.

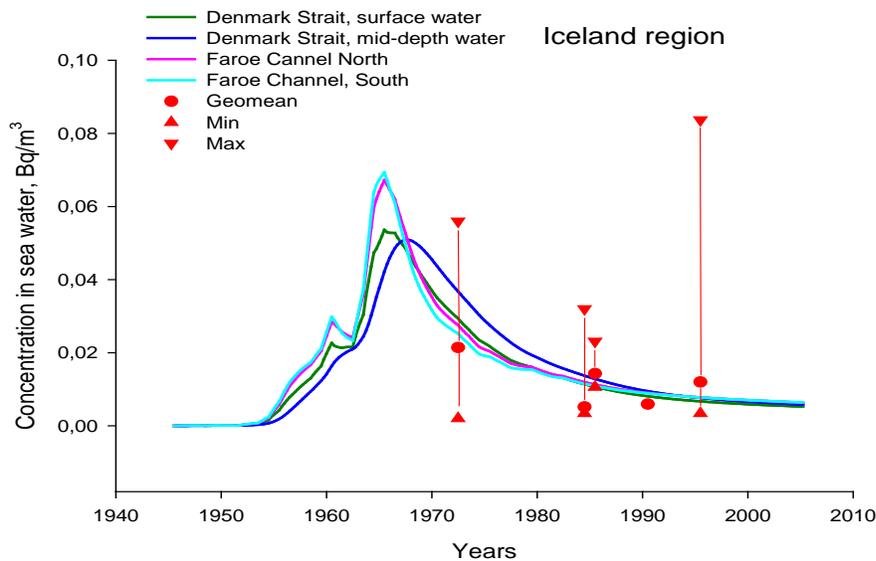
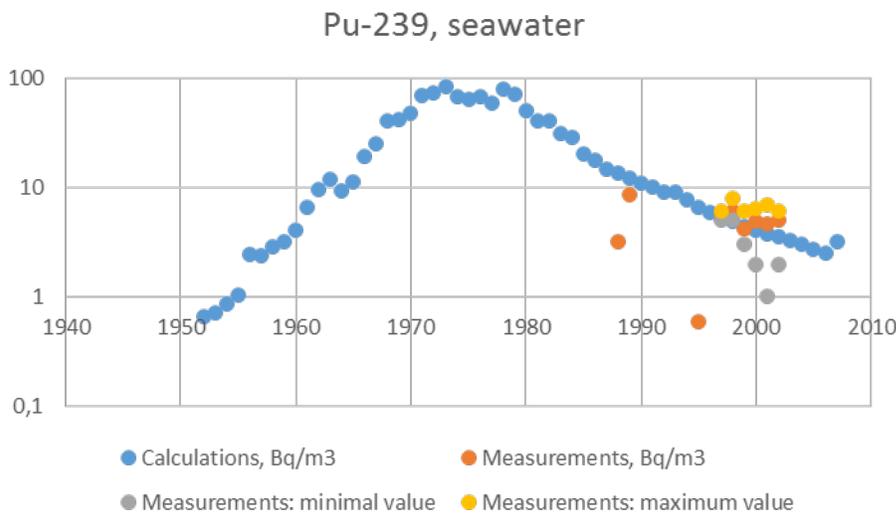


Figure 8.2-4: Filtered waters in the compartments: 128 DS1 –Denmark Strait Surface, 130 DS3 Denmark Strait Middepth, 134 FC1-Faroe Channel North, 140 FC7 Faroe Channel South.

8.2.1.2 Results of the AMIS model compared with measured data

Figures 8.2.5 and 8.2.6 show a reasonable correspondence between the results of the simulations and the experimental data. The environmental parameters of the model are therefore deemed suitable.



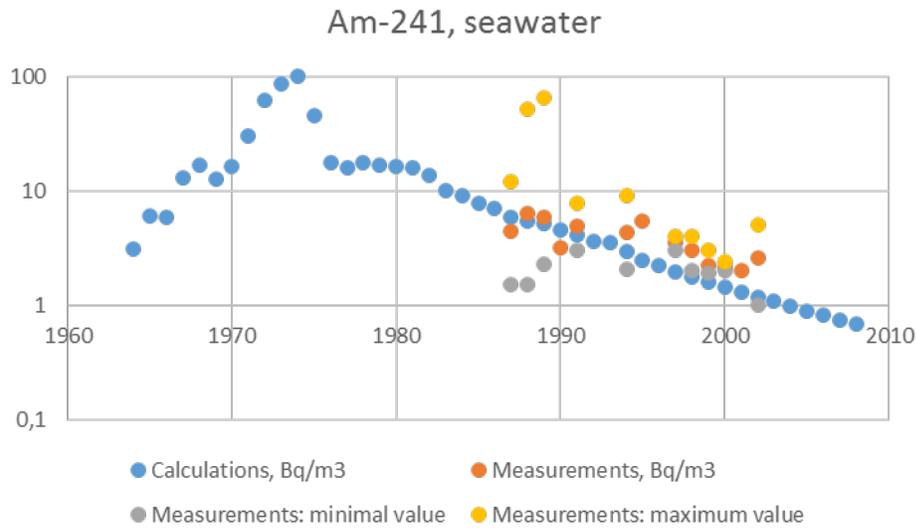
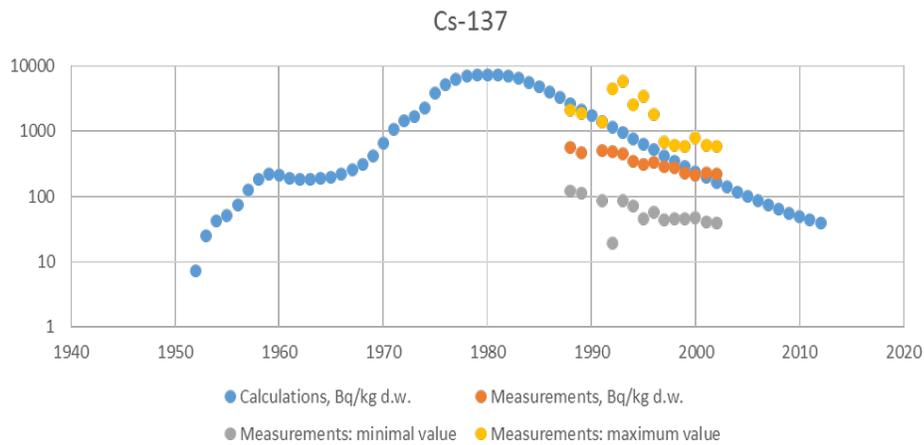


Figure 8.2-5: Comparison of the simulation results with experimental data in filtered water for Pu-239 (top) and Am-241 (bottom)



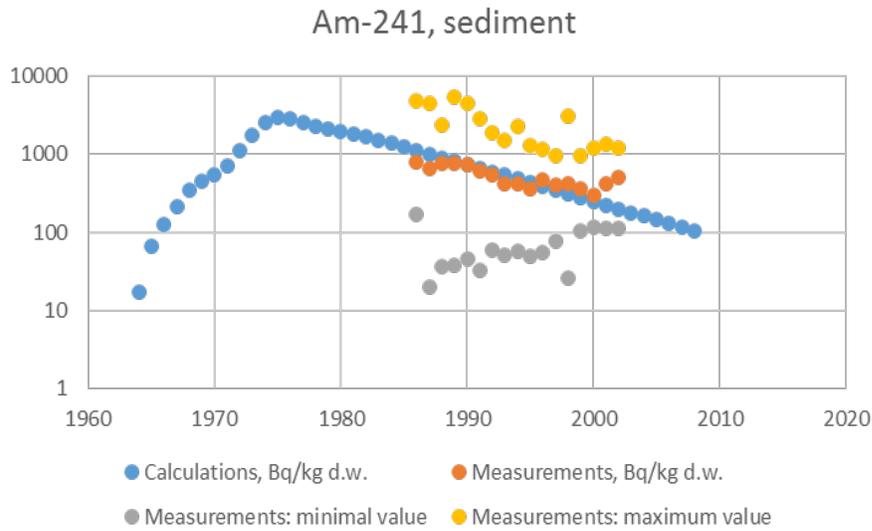
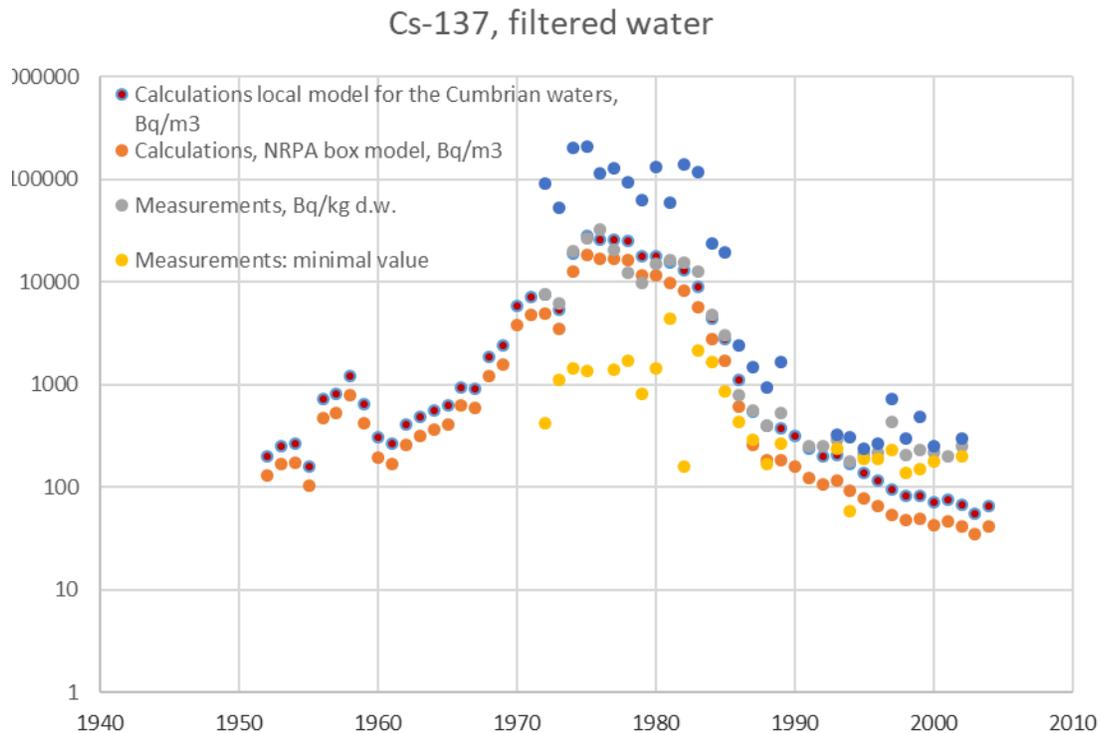


Figure 8.2-6: Comparison of the simulation results with experimental data in sediment for Cs-137 (top) and Am-241 (bottom)

8.2.1.3 Quantifying improvement the AMIS model versus the ARTICMAR model

Figure 8.2-7 shows a comparison of both models with experimental data for ¹³⁷Cs in the Cumbrian waters (the initial compartment for the releases from the Sellafield facilities). The releases corresponds to liquid discharges of ¹³⁷Cs from 1952 to 2016 years described in reports of British Nuclear Fuels (BNFL, 1976-2016), Radioactivity in Food and the Environment (RIFE, 1995-2014) and in MARINA II project (EC, 2003).



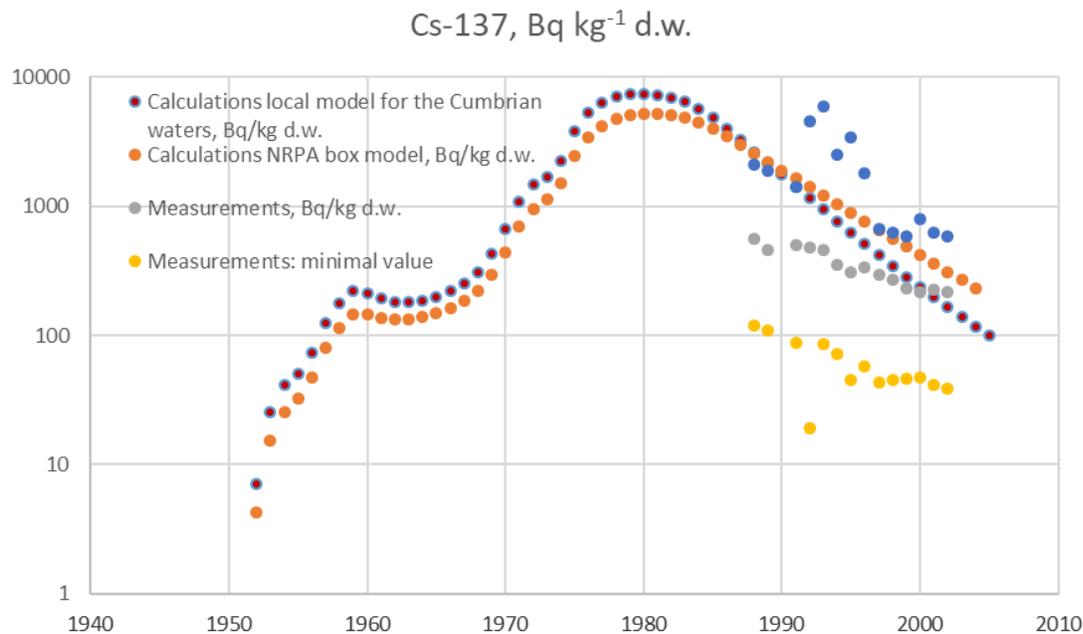


Figure 8.2-7: Comparison of simulation results with experimental data in sediment for the NRPA box model (ARCTICMAR) and for the regional model for the Cumbrian waters (AMIS) (in the blue points – measured max values)

Quantifying improvement the local model (AMIS) versus the NRPA box model (ARTICMAR) can be evaluated by the sum of the squares of the deviations between the calculation results and the experimental data.

Figure 8.2-7 as well as Table 8.2-1 indicates that use of the new sets of parameters leads to more precise calculations, which have been provided by the AMIS model.

Table 8.2-1: The sum of the squares of the deviations between the calculation results and the experimental data for ¹³⁷Cs

Water		Sediment	
ARCTICMAR	AMIS	ARCTICMAR	AMIS
5.61E+08	3.32E+08	1.11E+07	9.05E+06

8.2.2 AMIS model to quantify the fate of “hot particles” after long-term exposure in the Sellafield intertidal beach region

Assumptions and peculiarities of the release scenarios for “hot particles”.

Discharges of “hot particles” from the Sellafield nuclear facilities are not well understood. According to the recently published Sellafield Ltd report (2018), the non-numerical sources for the particle discharges can be described as (i) discharges via pipeline (for alpha and beta rich particles up to 1983 and 1985,

respectively) and (ii) due to debris from the Sealine Recovery Project 1990s and 2003-2006 (for all types of particles).

Based on this information the following scenarios have been chosen:

Scenarios for Cs-137:

Sc_c1L: $\varphi_{1L}^{(c)} \cdot Q/y$, where Q corresponds to liquid discharges from 1952 to 1985

Sc_c1: $\varphi_1^{(c)} \cdot 1\text{TBq}$ from 1952 to 1985

Sc_c2: $\varphi_2^{(c)} \cdot 1\text{TBq}$ from 1990 to 1995

Sc_c3: $\varphi_3^{(c)} \cdot 1\text{TBq}$ from 2003 to 2006

Sc_c4: $\varphi_{41}^{(c)} \cdot \text{Sc}_c1 + \varphi_{42}^{(c)} \cdot \text{Sc}_c2 + \varphi_{43}^{(c)} \cdot \text{Sc}_c3$

Sc_c5: $\varphi_{51}^{(c)} \cdot \text{Sc}_c1 + \varphi_{52}^{(c)} \cdot \text{Sc}_c2 + \varphi_{53}^{(c)} \cdot \text{Sc}_c3$

Scenarios for Am-241

Sc_a1L: $\varphi_{1L}^{(a)} \cdot Q/y$, where Q corresponds to liquid discharges from 1964 to 1983

Sc_a1: $\varphi_1^{(a)} \cdot 1\text{TBq}$ from 1964 to 1983

Sc_a2: $\varphi_2^{(a)} \cdot 1\text{TBq}$ from 1990 to 1995

Sc_a3: $\varphi_3^{(a)} \cdot 1\text{TBq}$ from 2003 to 2006

Sc_a4: $\varphi_{41}^{(a)} \cdot \text{Sa}_a1 + \varphi_{42}^{(a)} \cdot \text{Sa}_a2 + \varphi_{43}^{(a)} \cdot \text{Sa}_a3$

Sc_a5: $\varphi_{51}^{(a)} \cdot \text{Sa}_a1 + \varphi_{52}^{(a)} \cdot \text{Sa}_a2 + \varphi_{53}^{(a)} \cdot \text{Sa}_a3$

Scenarios Sc_c4, Sc_c5, Sc_a4 and Sa_a5 are prepared for the best comparison with experimental data after Monte-Carlo simulations for parameters $\varphi_{41}^{(c)}, \varphi_{42}^{(c)}, \varphi_{43}^{(c)}, \varphi_{51}^{(c)}, \varphi_{52}^{(c)}, \varphi_{53}^{(c)}, \varphi_{41}^{(a)}, \varphi_{42}^{(a)}, \varphi_{43}^{(a)}, \varphi_{51}^{(a)}, \varphi_{52}^{(a)}, \varphi_{53}^{(a)}$.

Further, the present modelling approach implements the non-instantaneous mixing of radioactivity in the model compartments. According to the algorithm of the simulations, "hot particles", released into the initial box (the Cumbrian Waters) via the pipeline, can reach the box boundaries within approximately 500 hours if they remain in suspension. However, a rough estimate of the time taken for the "hot particles" to fall to the seabed can be found by rewriting the Stokes law (Lamb, 1994) for the small spherical particles in a fluid medium:

$$t = \frac{4.5L\eta}{(\rho - \rho_w)gr^2} \quad (8.1)$$

where r is a radius of the spherical particle, which will reach the bottom in time t ; L is the distance to the bottom; η is the dynamic viscosity of seawater; ρ and ρ_w are the density of the particle and seawater, respectively, and g is the gravitational constant.

Complete information about dimensions of the "hot particles" is not available. Nevertheless, according to

the Sellafield Ltd (2017) report, almost 19% of “hot particles” have diameters more than 2 mm (such particles are classified as “objects”). It is easy to find from expression (3) that for relatively large particles with diameters 0.5 – 1 mm, residence time in the water column will be 0.4 – 1.6 seconds. Even for “very fine sand” (Krumbein & Aberdeen, 1937) with diameters 62.5–125 µm, this residence time will be approximately 7 – 27 min. Only particles with diameters less than 0.9 µm will have a residence time in the water column for more than 500 hours.

Based on the present assessment and acknowledging the approximations used in this approach, an assumption that almost all particles will reach the bottom sediment before they can reach the boundaries of the Cumbrian waters box seems reasonable. As such, the assumption of particles releasing directly into the sediment will be used as the release scenario for the “hot particles”.

Further, unlike the usual process of absorption and desorption of radioactivity on the surface of the suspended sediment particles in the water column, “hot particles” may include radioactivity within their bodies as soon as they come into the marine environment. Therefore, “apparent” k_d for the “hot particles” can be much higher than similar k_d for the usual particles suspended in the water column and sediment. For the following calculations values of $2 \cdot 10^4 \text{ m}^3\text{t}^{-1}$ and $5 \cdot 10^6 \text{ m}^3\text{t}^{-1}$ have selected for Cs-137 and Am-241, respectively. It is important to note that the set of other environmental parameters \vec{a}_μ is the same as those selected based on the liquid discharges.

8.2.2.1 Results for “hot particles” activity in the intertidal beach region

Calculations of hot particle activities in the intertidal beach region, where the monitoring program has been conducted, use the following expression:

$$A_S^{(MR)} = A_S^{(CW)} \frac{S_{MR}}{S_{CW}} P, \quad (8.2)$$

where $A_S^{(MR)}$ and $A_S^{(CW)}$ are predicted activities of the “hot particles” in sediment in the monitoring regions and in the Cumbrian Waters, respectively; S_{MR} and S_{CW} are areas of the monitoring regions and the Cumbrian Waters compartment, respectively; P is the probability that all particles in the monitoring regions have been detected (It is assumed that $P=1$ for the present calculations).

Figures 8.2-8 - 8.2-10 show simulation results for the beach region of the “Cumbrian Waters” sediment compartment. Results corresponded to the best fit for parameters $\varphi_{1L}^{(c)}$ and $\varphi_1^{(c)}$ for scenarios Sc_c1L and Sc_c1. Experimental data for the “hot particles” have been prepared based on information about the monitoring programme, published by Sellafield Ltd (2011-2017). Figure 8.2-8 shows scenarios Sc_c1L and Sc_c1. Figure 8.2-8 - 8.2-10 show comparison of predicted Cs-137 activity with experimental data for Sc_c1L and Sc_c1 correspondently.

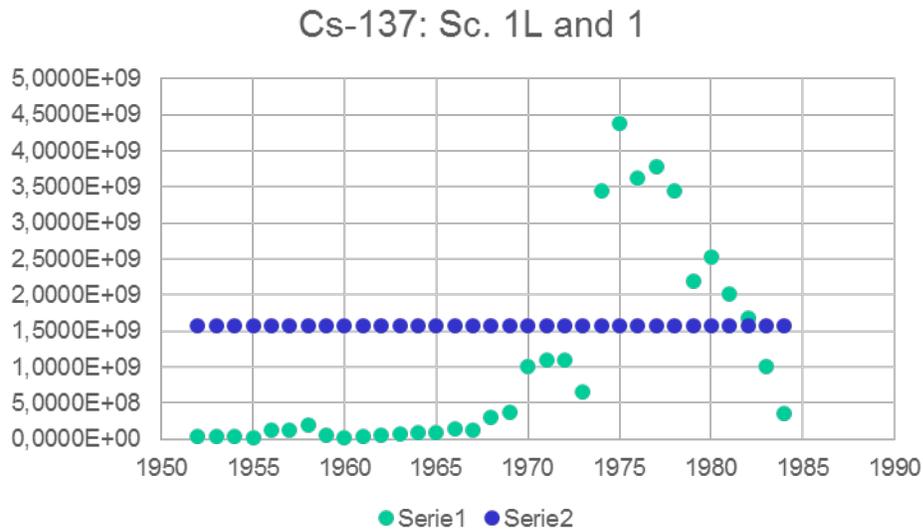


Figure 8.2-8: Scenarios Sc_c1L and Sc_c1. Activities of Cs-137 are shown in Bq.

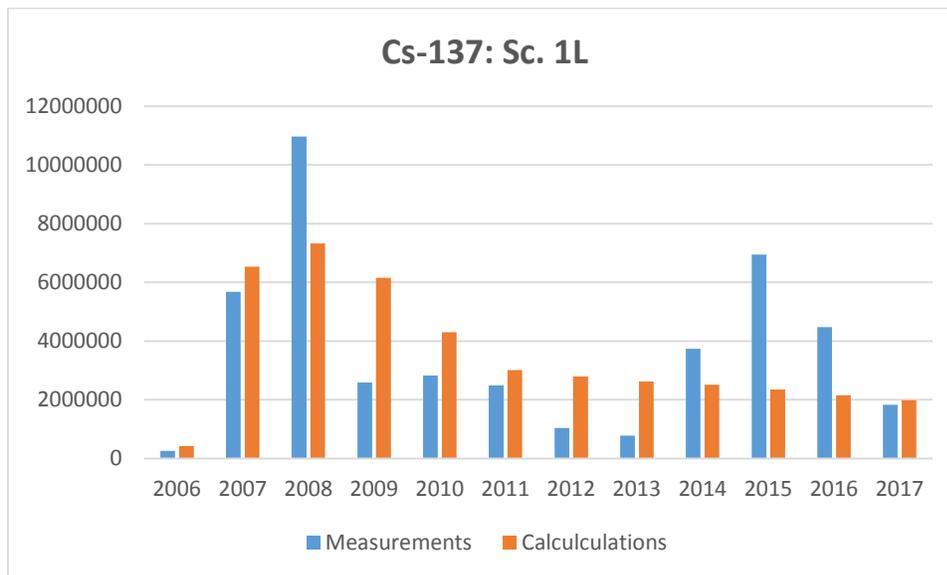


Figure 8.2-9: Comparison of the simulation results with the monitoring data for the sediment of the intertidal beach region according to scenario Sc_c1L. Activities of Cs-137 are shown in Bq.

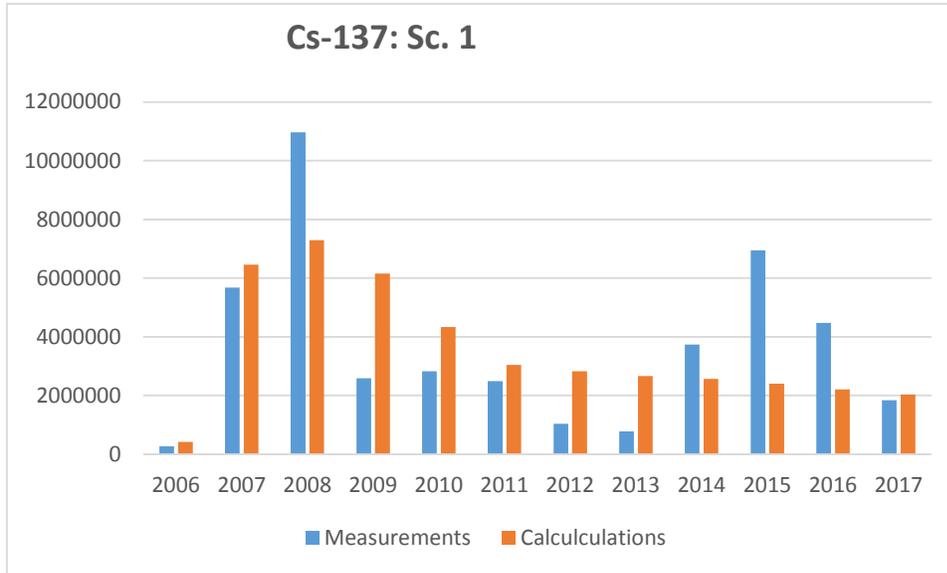


Figure 8.2-10: Comparison of the simulation results with the monitoring data for the sediment of the intertidal beach region according to scenario Sc_c1. Activities of Cs-137 are shown in Bq.

Figures 8.2-11 and 8.2-12 show similar results for scenarios Sc_a1, Sc_a2 and Sc_a3 for Am-241.

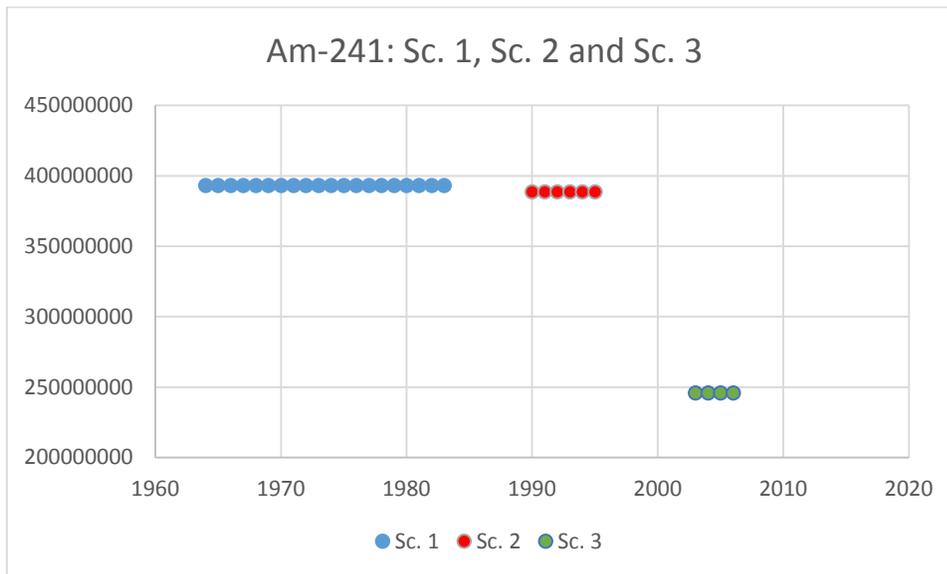


Figure 8.2-11: Scenarios Sc_a1 and Sc_a4. Activities of Am-241 are shown in Bq.

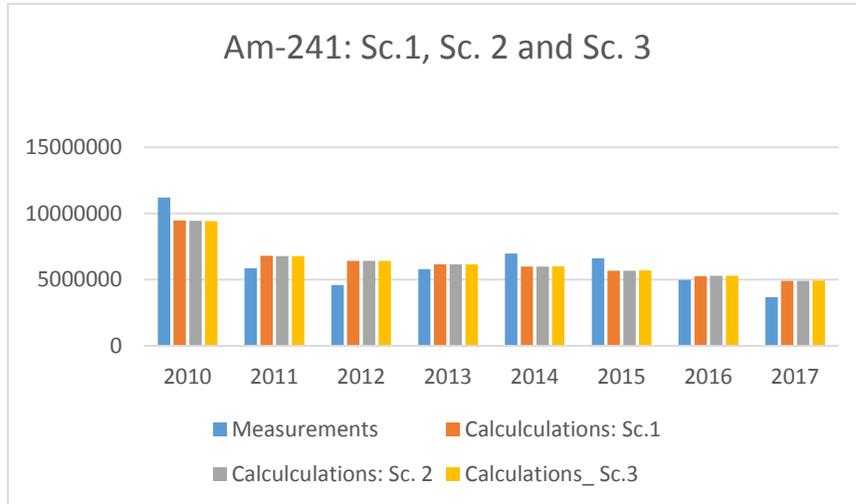


Figure 8.2-12: Comparison of the results of simulations with monitoring data for the sediment of the intertidal beach region according to scenario Sc_a1, Sc_a2 and Sc_a3. Activities of Am-241 are shown in Bq.

Figures 8.2-9, 8.2-10 and 8.2-12 show a reasonable agreement between experimental data.

It is important to note that according to the Sellafield Ltd (2017) report, at the end of 2009 the monitoring process of “hot particles” was improved, in part with regards to the detection of particles containing Am-241. This information has cast some doubt on the possibility to assume that $P=1$ in Figure 8.2-12 before 2010. Therefore, Figure 8.2-12 shows comparisons between prediction and monitoring data for Am-241, but only for experimental data from the monitoring program after 2009.

Similar calculations have been executed for all scenarios. The best comparison with experimental data for Cs-137 corresponds to scenario Sc_c5 with the total released activity of $1.6 \cdot 10^{10}$ Bq. Similarly, the best results for Am-241 corresponds to scenarios Sc_a1 with the total released activity of $8.3 \cdot 10^9$ Bq

8.2.2.2 Estimation of uncertainties

Calculations have demonstrated that very different scenarios can provide similar results because of the complexity of the model / system (Iosjpe, 2011B; 2014). In such conditions, additional information about scenarios seems necessary.

It is possible to estimate uncertainties by the global sensitivity index (Till and Meyer, 1983)

$$S^{(G)}(P) = 1 - \frac{P_{min}^{(S)}}{P_{max}^{(S)}}, \quad (8.3)$$

where $P_{min}^{(S)}$ and $P_{max}^{(S)}$ are minimum and maximum absolute values of the state parameter $P^{(S)}$ within the range of parameter P , which have been described in expression (8.3). According to construction, $S^{(G)} = 0$, when there is no influence of the evaluated parameters to the state parameter. $S^{(G)} = 1$, when evaluated parameter has strongest influence to the state parameter.

$P_{min}^{(S)}$ and $P_{max}^{(S)}$ can be estimated with help of the local sensitivity indexes. One example for the Am-241 hot particles is shown in Figure 8.2-13.

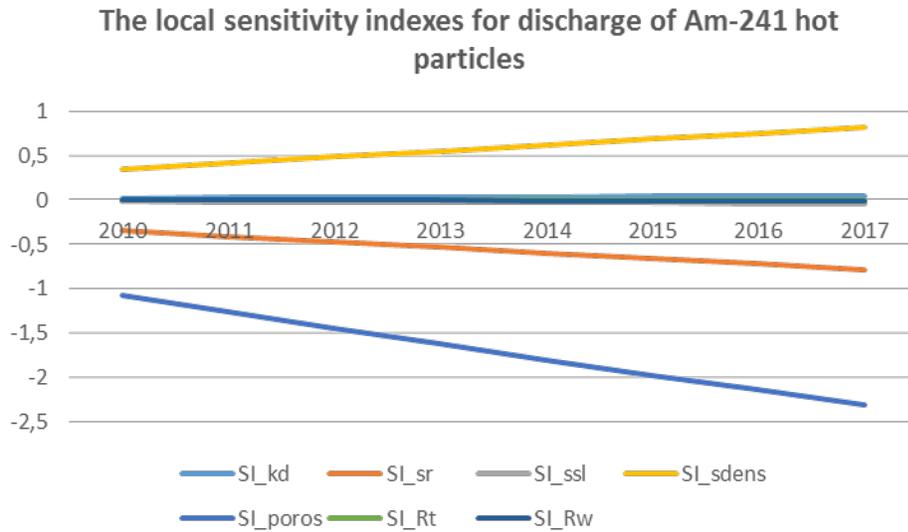


Figure 8.2-13: Influence of the model parameters to average concentration per year for Am-241.

Further, the most suitable set of the environmental model parameters has been selected. Therefore, the range for the parameters has been chosen to be 10%, similar to (Nielsen, 1998). The range for the apparent sediment distribution coefficient has been chosen from standard k_{d0} by to $k_{d0} \cdot 10$ according to recommendation from IAEA (2004).

Calculations indicates that for Cs-137

$$S_p^{(G)} = 0.928,$$

where the global sensitivity index $S_p^{(G)}$ is associated with parameters uncertainties for the best release scenario for Cs-137.

The global sensitivity index

$$S_{sc}^{(G)} = 0,987$$

is associated with scenarios uncertainties.

The estimation for the total release for the Cs-137 can be described as $6.9 \cdot 10^8 < 1.6 \cdot 10^{10} < 7.6 \cdot 10^{10}$ Bq.

The global sensitivity index associated with parameters uncertainties for the best release scenario for and scenarios uncertainties Cs-137.

$$S^{(G)} = 0,991$$

Similar, the global sensitivity index $S^{(G)}_p$ is associated with parameters uncertainties for the best release scenario for Am-241

$$S^{(G)}_p = 0.921,$$

while the index

$$S^{(G)}_{sc} = 0,897$$

is associated with scenarios uncertainties.

The estimation for the total release for the Am-241 can be described as $9.8 \cdot 10^8 < 8.3 \cdot 10^9 < 4.4 \cdot 10^{10}$ Bq.

The global sensitivity index associated with parameters uncertainties for the best release scenario for and scenarios uncertainties Am-241.

$$S^{(G)} = 0,978$$

These results show that to assess the overall activity for the best release scenarios it is important to assess the uncertainties associated with both the model parameters and the choice of a possible release scenario.

So, a specific regional model for the Cumbrian Waters (AMIS) has been constructed to investigate the fate of radionuclides (especially those associated with “hot particles”) in the intertidal beach region near the Sellafield nuclear facilities. Corroboration of the model has been provided based on the comparison of the results of the calculations with experimental data for the liquid discharges of Cs-137, Pu-239 and Am-241 into the Cumbrian waters.

Assumptions for the release scenarios for Cs-137 and Am-241 associated with “hot particles” have been proposed. The comparison of the results of the calculations with monitoring data for “hot particles” demonstrates that the present modelling approach for the constructed regional model can be used to quantify the fate of the “hot particles” after long-term exposure in the intertidal beach region.

Sensitivity analysis has been used for the evaluation of the influence of the model parameters to the end points of modelling and selection of the suitable environmental parameters.

The uncertainty analysis was used to estimate the released total activity of the Cs-137 and Am-241 hot particles. The results show that to assess the overall activity for the best release scenarios it is important to assess the uncertainties associated with both the model parameters and the choice of a possible release scenario.

9 DISCUSSION

For the Norwegian Fen forest site, co-located measurements of gamma air kerma rates and activity concentrations in soil have been used to assess the performance of three external exposure models as implemented in CROM, NORMALYSA and GRANIS. The level of complexity of the three models varies with input requirements encompassing soil activity concentrations, radionuclide, soil composition, soil porosity, soil density and depth of contamination. Each model has been used to calculate gamma air kerma rates at 1 m above the ground using measurements of activity concentrations in soil and these have been compared with co-located measured air kerma rates. A statistical analysis of the results has been carried out for each model using the guidance provided in Section 3.4 - 3.6. Initially, the root mean squared logarithmic error (RMSLE) was the primary indicator used to provide a measure of model performance. The results of this analysis suggest that the best performing model is NORMALYSA followed by CROM and then GRANIS i.e. the most detailed model is performing least well. Possible reasons for this have been discussed and it can be concluded that a model should not be treated as a black box but knowledge of the calculation methods used is necessary if a more fundamental understanding of model performance is to be achieved. In addition, it is important to use appropriate validation datasets. In this study, the data were very heterogeneous and in an ideal situation a larger number of co-located measurements would be used to validate the models.

For the Belgian NORM forest site, FORESTCROM which apply a simple dynamic model based on published transfer factors was applied together with two water flow models (TALAL and ECOFOR), developed specifically for this site, that accept multiple soil layers of different physical and hydraulic characteristics. TALAL uses the full representation (coupling Richards equation and the advection dispersion equation) and ECOFOR a more simplified approach, as described in the model's section. TALAL and ECOFOR models are driven by evapotranspiration, which is automatically calculated from meteorological data in case no values are directly available to the codes as input, and were fully coded and verified (not validated). All the models were parameterised with data from measurements produced under TERRITORIES project, complemented with literature values. All the three models perform very well for the Belgian NORM case, although it must be pointed out that all of them were calibrated for the site. The simulation results from ECOFOR and TALAL models reflect the nature of the study region. ECOFOR's calculations provide the closer match for the ^{238}U , ^{210}Pb and ^{226}Ra soil/tree transfer factor, compared with the values calculated from measured activity concentrations in the sludge and root samples. The transfer factors predicted using the TALAL model are also in excellent agreement with the measured values when adjusting the absorbing power of the tree root system.

For the Fukushima forest site three models were applied: FORESTCROM, a very simple dynamic forest model developed for TERRITORIES; TREE4, a dynamic process-based forest model developed at IRSN; and an improved version of TREE4 where major improvements were achieved thanks to the knowledge gained from field observations acquired over the period 2011-2017 in Fukushima coniferous forests. All the three models achieve good performance against the measurements carried out in the Japanese forests; in view of the results of the application of the performance metrics (see appendices). FORESTCROM result in a good performance for the intercepted Cs in the canopy and in the final transfer of the Cs to the wood, once local parameters (mainly the interception factor) were calibrated. However, in the case of the concentration in the soil the model results in an important underestimation. Both TREE4 versions result in better results in general, although for the advanced version (TREE4-advanced) the accuracy was improved in comparison with the old TREE4 model for both Cs inventories and depuration fluxes, due to a better

quantification of the throughfall contribution, which itself indirectly results from the introduction of the foliage excretion process. Predictions of Cs concentrations in tree organs were also improved especially in foliage and trunk wood. This could be attributed to: (i) the differentiation of the canopy into foliage and branch compartments with distinct eco-physiological characteristics, (ii) the influence of foliage excretion in the long-term decrease of foliage contamination and (iii) the contribution of the foliage-to-wood translocation process to wood contamination. Conversely, the predicted concentrations in the upper organic and top 20-cm mineral layers, using the TREE4 model, could not be improved. Nonetheless, unlike the simple version, the “advanced” model can now potentially predict the vertical distribution of Cs within the mineral profile, this information being required when evaluating ambient dose rates in contaminated forests. Although not shown in this document, we further demonstrated that the model’s accuracy was significantly improved, especially in tree vegetation. We also demonstrated that the observed (spatio-temporal) variability remained strongly under-predicted by the model. This was attributed to the absence of seasonality in the model and the non-consideration of model/scenario uncertainties in the analysis.

Finally, for the marine Sellafield site, a specific regional model for the Cumbrian Waters (AMIS) has been constructed to investigate the fate of radionuclides (especially those associated with “hot particles”) in the intertidal beach region near the Sellafield nuclear facilities. The methodology is based on the earlier ARCTICMAR model. Corroboration of both the ARCTICMAR and AMIS models has been provided based on the comparison of the results of the calculations with experimental data for the liquid discharges of radionuclides into Cumbrian waters. Assumptions for the release scenarios for Cs-137 and Am-241 associated with “hot particles” have been proposed for the AMIS model. The comparison of the results of the calculations of the total activity associated with “hot particles” with monitoring data for “hot particles” demonstrates that the present modelling approach for the constructed regional model can be used to quantify the fate of the “hot particles” after long-term exposure in the intertidal beach region.

The analysis of data for all case studies will be returned to in Section 9.2 where the quantitative and qualitative performance assessment for each of the models will be considered systematically. Before this is covered, however, we consider it necessary to present what we determine to be good modelling practice by considering tiered approaches to assessment, the common steps required in modelling and quality assurance issues.

9.1 Good modelling practice

9.1.1 A tiered approach to assessments

There is a general requirement in the field of radiological protection to ensure that the application of the system of protection is commensurate with the radiation risks associated with the exposure situation (ICRP, 2006; IAEA, 2018). This requirement extends to the level of detail and resources required in the assessment of doses associated with a planned exposure (IAEA 2018, IAEA, 2001). The standard approach for a radiological environmental risk assessment is to implement a tiered or graded approach. This can be evidenced in the USDoE’s graded approach (USDoE, 2002) and the ERICA integrated approach (Larsson, 2008; Brown *et al.*, 2016) which both adopt this methodology in determining the risk of harm to non-human biota arising from radionuclides released to or present in the environment.

Such an approach essentially attributes the level of detail and resources required for the assessment to the perceived risk associated with the hazard. Typically the assessment would begin with a lower tier,

screening assessment, where conservative assumptions are adopted (using simple models with generic parameters) to obtain pessimistic predictions of the radiological impacts. If at this stage it can be demonstrated that the risks are trivial then there may be no need to perform more detailed calculations. However, where this is not the case, the assessment process should move to higher tiers where more site specific details (and concomitantly more resources) are required to provide an estimate of the radiological impact. For the higher tiers, the additional resources might be spent on more detailed model parameterisation, more comprehensive data collation or model development.

At each stage of the tiered assessment, it is necessary to demonstrate that the model being used is fit-for-purpose. At the initial screening stage one of the criteria would be that the model needs to be cautious i.e. predictions would not expect to be exceeded under any foreseeable circumstances. For higher tiers of assessment, it would be necessary to increase resources appropriately to make use of site-specific data and bespoke models to improve the predictions. This process should continue until it can be demonstrated that risks are acceptable, taking into account associated uncertainties, or unacceptable in which case the exposure scenario cannot be permitted to occur.

The tiered approach is arguably only relevant in the context of assessments associated with regulation where the endpoint is to determine whether the risk associated with an exposure is acceptable. For other considerations, for instance the application of models for research objectives where a deeper understanding of processes is desirable, the step-wise approach is unlikely to be suitable.

9.1.2 The steps in model development and implementation

Refsgaard *et al.* (2007) provided an outline of the typical steps associated with a model development and implementation procedure and how these are normally linked to administrative/decision-making processes. Although the outline is specific to water resources management *per se*, it remains highly pertinent to environmental management in general and thus has relevance for impact assessments involving the modelling of radioactivity behaviour and fate. The five steps considered by Refsgaard *et al.* (2007), which may be considered as guidance with regards to an ideal sequence involving model development and implementation, are : (1) Model Study Plan, (2) Data and conceptualisation, (3) Model set-up, (4) Calibration and validation and (5) Simulation and evaluation.

A more detailed description regarding what these steps involve and guidance on how each of these points may be applied in practice, drawing on the experience gained from model application to the study sites in TERRITORIES WP1 as described in previous chapters of this report, is given below.

9.1.2.1 Model Study Plan

This initial step is primarily linked to formulating the problem and defining the assessment context. It would be typical to ask various questions in relation to the purpose of the assessment and modelling requirements. As exemplified with the work performed in relation to environmental impacts from ionising radiation (Larsson, 2008) consideration might be given to ecological, political and societal issues when deciding on procedures and methods, who to involve, and any benchmarks or assessment criteria that the outcome will be compared to. The availability of relevant data for conducting an assessment might also be considered at this stage. An important, but often ignored task is the analysis and determination of the

various requirements associated with the modelling study in terms of the expected accuracy of modelling results. The acceptable level of accuracy will vary from case to case and might be considered within a socio-economic context. Without defining this at the start of the process, establishing whether a model is fit-for-purpose and whether the appropriate level of complexity has been selected becomes challenging. A **Model Study Plan should always constitute an initial component of a process to establish the fit-for-purpose level of complexity required for a model by, inter alia, clearly defining the criteria against which the model will be judged.**

9.1.2.2 Data and conceptualisation

This step could either involve the development of a conceptual model or the consideration of existing models/computer codes in view of their applicability to the case in hand. Ideally, insights should be gained into the main processes and interactions in the system evaluating whether these can be modelled in sufficient detail to meet the requirements specified in the model study plan. One way to achieve this may be through the application of a Features-Event Processes (FEPs) analysis as described in Section 3.1. An overview and guidance as to how best practice methods including FEPS are applied in the context of modelling radioactive contaminations in forests are discussed by Diener *et al.* (2017). These authors go on to describe how FEPs may be presented as an interaction matrix thus forming the basis of a mathematical representation of the conceptual model. They note the importance of mathematically representing processes to an adequate or fit-for-purpose level of complexity and suggest that one way of achieving this may be keeping the number of independent model parameters as low as practicable (whilst still maintaining model efficacy). This would have the effect of reducing the overall model output uncertainty. This builds upon the guidance of Kirchner and Steiner (2008) who noted that in cases where various models are available for a given radioecological problem the model with the optimum structure should be selected. In order to identify this, the following criteria were considered helpful:

- The dominant transport and transfer processes should be taken into account with minimum model complexity.
- The number of model parameters for which values have to be specified should be reasonably small.
- The model parameters should ideally be only moderately sensitive and moderately variable/uncertain.
- Model parameters should be constant and not represent hidden dynamic processes.

In addition to this, and as touched upon in Section 3.4, there are other methods such as Akaike Information Criterion (AIC) (Burnham and Anderson, 2002), that can be used to compare models to identify which has the optimum structure. For example, AIC deals with the trade-off between the goodness of fit (between model predictions and empirical data) and the simplicity of the model. In this way, models are penalised for over-parameterisation and AIC provides a relative measure for a given model compared to others of the degree of overfitting and underfitting of datasets. **The recommendation is that the above mentioned criteria, as defined by Kirchner and Steiner (2008), and the tools described in Section 3.4 - 3.6, should be used in establishing which model provides the most suitable fit-for-purpose level of complexity if several models be available.**

Further considerations might be afforded the spatial and temporal detail required, the system dynamics, the boundary conditions and specifications as to how the model parameters can be determined from available data. In order to be able to assess uncertainty in model structure, the need to model certain processes in alternative ways or to differing levels of detail should be evaluated. Data collations should be made with specific applications in mind. For the case of the TERRITORIES project, the TLD has been collated primarily with a view to model testing, for the purpose of demonstration, although other goals for the development of a dataset, such as model parameterisation, calibration and validation, could be deemed appropriate.

9.1.2.3 Model set-up

The step of model set-up either involves the transformation of the conceptual model into a mathematical representation and computer coding of this representation or the adoption and configuration of existing models. The computer coding can take various forms from fundamental applications (programming using computer languages such as Java, C++, Fortran etc.) to the utility of 'modelling platforms' such as Model-maker, MATLAB, R, Gold-sim, ECOLEGO (e.g. Børretzen & Salbu, 2000, Avila *et al.*, 2005, Lee & Kim, 2017).

In the TERRITORIES project application, various categories of the abovementioned methods have been employed in fulfilling the model set-up stage. For example, application of the modelling platform Model-maker has been utilised in the development of ECOFOR for the Belgian NORM site (Section 6.2.2) whereas a re-programming of an original model code (Fortran/time-zero) has been required in the adaptation of the DSA-marine-Box model to a model that can account for the behaviour of particles from the Sellafield site (Section 8.2.2).

Data processing in order to prepare various data files in a format necessary for executing the model is often a challenging and time consuming task. To illustrate this point, the work conducted in applying various 'off-the shelf' models to the Norwegian NORM case (Fen Field), as described in Sections 5.2.2 and 5.2.3, illustrated that the modeller can sometimes be confronted by a lack of compatibility between available datasets and the requirements of the model. This, for the given example, in one case (Section 5.2.2) involved a lack of spatial co-locality between input data (activity concentrations in soil) and outputs (activity concentrations in biota) and in another case poor specificity in relation to radionuclide distributions with depth (Sections 5.2.3) rendering model air dose rate determinations, based on soil radionuclide concentrations, uncertain. There is arguably no definitive guidance that can be provided in addressing such challenges other than **recommending that the modeller should be explicit in terms of the assumptions made when rendering available datasets compatible as inputs to the model and should provide a range of plausible values for the dataset in an attempt to account for uncertainty.**

9.1.2.4 Calibration and validation

Calibration may involve adjustment of the model parameters within the bounds of the uncertainties to obtain a model representation of the processes of interest that satisfies pre-agreed criteria such as Goodness-of-Fit. As noted by Arnold *et al.* (2012), calibration is an effort to better parameterize a model to a given set of local conditions, thereby reducing the prediction uncertainty. This may be performed by carefully selecting values for model input parameters (within their respective uncertainty ranges) by

comparing model predictions (output) for a given set of assumed conditions with observed data for the same conditions.

Calibration has been elegantly demonstrated in the re-estimation of parameters in the forest stand module, such as those involved in the allometric and growth curves, of the TREE-4 Model (Section 4.2.4) and the derivation of radionuclide sorption rate for given soils derived via calibration trials in the case of the ECOFOR hydrological sub-model (Section 6.2.2)

Model validation is typically undertaken through the comparison of model output with independently measured values. A methodology explaining how this might be achieved in practice is provided in Section 3.4 - 3.6 of this report. To reiterate, the principal numerical metric that have been recommended for use in providing a measure of how well model outputs correspond to empirical data is the Root Mean Squared Logarithmic Error (RMSLE) owing to the fact that the method deals appropriately with highly variable data. Other metrics may also be applied, as explained in this section of the report, to obtain quantitative information regarding the capability of the model to avoid overestimation or underestimation of the data and to preserve data pattern. By way of example, RMSLE was applied in comparing the performances of the models CROM, NORMALYSA and GRANIS to the Norwegian Fen Forest site, specifically for the case of deriving Gamma air kerma rates from inputted data on NORM radionuclide activity concentrations in soils (Section 5.2.3). The general view derived from this experience was that the RMSLE has a high practicability allowing the efficacy of models, at least in the given examples, to be compared in a reasonably straightforward manner. Nonetheless, the caveat should be added that great care should be afforded the alignment of model output and empirical datasets when deriving such metrics. This point may seem trivial, but even for the relatively simple comparison of modelled versus empirical data pertaining to gamma air kerma rates, account needed to be taken of the cosmic component of radiation, conversion factors (ambient dose to air kerma) among other considerations.

Admittedly, the application of the aforementioned methodology prepared for this report falls some way short of the level where it may be considered a full or comprehensive validation. For such a goal to be achieved a much more elaborated analysis would necessarily be invoked.

In addition to these metrics, a set of qualitative indicators is also presented in Section 3.6, for the purpose of judging the quality of a model if data are lacking. Both qualitative and quantitative indicators are merged into a performance table to provide a practical structured tool for dealing with the assessment of model performance.

Diener *et al.* (2017) opined that representative data are the basis for the development, calibration and validation of robust radioecological models. In this regard, Ramsey and Hewitt (2005) provide a suitable methodology for the evaluation of sample representativeness for a more general setting although other published information sources might also be consulted. It is recommended that the purpose of the sampling campaign and the expected variability should be taken into account (Diener *et al.*, 2017). For an example involving the calibration and validation of radioecological forest models, Diener *et al.* (2017) considered that the collation of extensive data sets characterising both the concentration levels of radionuclides in different media and the status of the forest ecosystem, might be necessary. If a radioecological model is designed to provide the variability of given endpoints, data for model validation have to be collected accordingly with due attention to factors such as sample size and representativeness. Furthermore, seasonal changes and specific climatic conditions often require frequently sampled data.

Diener *et al.* (2017) noted that a strong interaction between experimentalists and modellers would be advantageous to obtain well-documented high-quality data as the basis for model development, calibration and validation.

Finally, in line with Refsgaard *et al.* (2007), the reliability of model simulations for the intended type of application can be assessed through uncertainty analyses. This would imply that some form of probabilistic/stochastic modelling capability should be incorporated or developed within the approach. This could form a component of the model evaluation, as considered below in the next step, and further in the uncertainty assessment (Appendix 12.2).

The results should be described so that the scope of model use and its associated limitations are documented and made explicit. This is essentially where the fitness-for-purpose of the model may be established. The possibility exists that the model does not meet the requirements set out in the model study plan, for example, the model may not provide the (pre-defined) level of accuracy required or lack coverage of processes considered essential in characterising the system.

9.1.2.5 Simulation and evaluation

In this step, simulations are performed to meet the objectives and requirements of the model study. This last step has really fallen beyond the scope of the analyses made in this report but in the 'real world' it is reasonable to expect that simulation results may be used in subsequent decision making (e.g. for planning or design purposes) or to improve understanding (e.g. of the radionuclide transfer processes in the study area).

Mattot *et al.* (2009), albeit in relation to integrated environmental models, define seven subjective categories of methods for quantitative model evaluation namely : data analysis, identifiability analysis, parameter estimation, uncertainty analysis, sensitivity analysis, multi-model analysis, and Bayesian networks. Of particular note is Uncertainty Analysis (UA) methods, which are used to propagate sources of uncertainty through the model to generate statistical moments or probability distributions for various model outputs. A full description of uncertainty analyses has been described elsewhere in the TERRITORIES project (Urso *et al.*, 2019). Refsgaard *et al.* (2007) emphasise the point that it is important to carry out suitable uncertainty assessments of the model predictions in order to arrive at robust decisions. This is considered in more detail, with concomitant guidance, in Appendix 12.2.

Furthermore, the subject of uncertainty analysis has at least been touched upon within the deliberations provided in earlier parts of this report. In propagating uncertainties when performing simulations using the TREE-4 model, Monte-Carlo simulations (10000 runs) were undertaken taking into account parameter uncertainties, and correlations between them (Appendix 12.2). Providing an analysis in this way, led the modellers to conclude that predicted variability/uncertainty was too low for all radioactive quantities and provide reasons for this reflecting that adequate account may not been taken of infra-annual effects, scenario uncertainties, structural uncertainties and uncertainty associated with radionuclide deposition levels. It is difficult to draw more generic guidance from this quite specific example, other than to say that **efforts should be made to characterise the uncertainties and evaluate the efficacy of the model using some of the approaches/tools described elsewhere (Urso et al, 2019) in addition to those considered in Appendix 12.2.**

As with the other steps, Refsgaard *et al.* (2007) notes that the quality of the results needs to be assessed through internal and external reviews that also provide platforms for dialogues between the modeller, reviewer and, often, stakeholders/public. This recommendation is fully supported within this current guidance.

9.1.3 Quality Assurance

Quality assurance (QA) should form an integral part of any approach to establish whether a model is fit for purpose and has the correct level of complexity. It should go without saying that without establishing whether a model's outputs are robust and reliable, there is little that can be said about whether the model is suitable for solving a given problem.

Calder *et al.* (2018) offer a guide to the process of commissioning, developing and deploying models across a wide range of domains. In so doing, the authors provide two checklists to help potential modellers, commissioners and users ensure they have considered the most significant factors that will determine success. In posing some of the questions that need to be answered before and during the creation and use of a model, Calder *et al.* (2018) specify the following for QA:

- What verification procedures will be used to check that the model works as expected?
- How will the model be validated, and what data will be used for doing so?
- Is there a schedule of reviews to ensure that the model remains up to date?

Although HM Treasury (2013) provided a review specifically to QA of analytical models that inform government policy in the UK. Some of the observations and recommendations made in that report are pertinent to the broader issues of modelling QA for our purposes. HM Treasury (2013) considered that best practice in QA fell under two headings: modelling environment, and process. The former was defined as involving a culture where leaders value and recognise good QA. This would require adequate capacity, including specialist skills and sufficient time to conduct QA effectively. The process heading encompassed systematic approaches to make QA accessible, easy and comprehensive. This might require clear guidance on QA, and clear documentation for every model. A list of QA procedures was also described and mapped (in relation they applied to models underpinning UK policy) and included, *inter alia*, internal and external peer review, internal and external audit, Developer testing (use of a range of developer tools including parallel build and analytical review or sense check), Use of version control (use of unique identifier for different versions of a model), QA guidelines & checklists and Transparency (where the model is placed in a wider domain for scrutiny).

Many QA guidelines exist a useful overview for which is provided in Refsgaard *et al.* (2005) in the context of water management but for which a more generic modelling application is evident. A useful definition provide in this paper for QA is "the procedural and operational framework used by an organisation managing the modelling study to assure technically and scientifically adequate execution of all tasks included in the study, and to assure that all modelling-based analysis is reproducible and defensible. As noted by Refsgaard *et al.* (2007), QA improves the chances that best practise is used. Furthermore, the possibility to involve stakeholders is introduced into the modelling process in a formalised framework, and the transparency and reproducibility may be improved.

9.2 Performance assessment analysis of the models. Comparison of the models.

A comparison of all widely used and advanced model for three forest sites and one marine site was performed (see Tables 12.1-1 – 12.1.22 in Appendix 12.1).

Three indicators have been used for this comparison: RMSLE, BIAS and MG (RMSLE and MG are defined as unitless statistics, but BIAS has units, which according to Eq. (3.4) – (3.7) are the same as the initial values compared (see description in Section (3.4) for BIAS, MG; and section 3.6 for RMSLE).

In the Table 9.2-1, this information has been summarised. The quantitative and qualitative performance assessment of all 6 widely used and 5 more advanced models are presented together with sites where they were applied.

Table 9.2-1: Quantitative and qualitative performance assessment of models

N	Name of model	Quantitative indicators of model			Amount of measurements used	Magnitudes compared ^{3,4}	Qualitative indicators of model		Site where applied
		RMSLE ⁵	BIAS ⁶	MG ⁷			Number of parameters involved	Number of processes considered	
WIDELY APPLIED MODELS									
1	CROM	0.59	2.46	0.69	5	H*(10)	> 100	1	NORWEGIAN FEN FOREST SITE
2	NORMALYSA	0.23	-0.23	1.37	5	H*(10)	12	1	NORWEGIAN FEN FOREST SITE
3	FORESTCROM	0.0457	11.4	0.9	3	AL	11	4	BELGIAN NORM SITE
4	FORESTCROM	0.72	see appendix ⁸	2.8	15	AL	11	4	FUKUSHIMA
5	TREE4-simple	0.41 ⁹	see appendix	0.97 ⁹	83 ⁹	AL	22	14	FUKUSHIMA
6	ARCTICMAR	0.44	see appendix	1.35	131	AL	2000	24	SELLAFIELD
MORE ADVANCED MODELS									
1	GRANIS	0.759	0.68	2.95	5	H*(10)	10	4	NORWEGIAN FEN FOREST SITE
2	ECOFOR	0.015	0.0007	1.01	3	AL	61 general 115 site spec.	25	BELGIAN NORM SITE

³ H*(10) - ambient dose equivalent ($\mu\text{Gy h}^{-1}$)

⁴ AL -activity level (Bq/m^3 , Bq/kg , Bq/l , etc., depends on results generated by the model)

⁵ unitless statistics

⁶ the same units as the initial compared values

⁷ unitless statistics

⁸ Different ALs are modelled. More detailed information see in correspondent to each model appendix.

⁹ Values calculated for Activity Concentrations in soil layers and tree organs (ie activity inventories and fluxes are not accounted for)

3	TALAL	0.098	0.0024	1.04	3	AL	30	21	BELGIAN NORM SITE
4	TREE4-advanced	0.24 ⁹	see appendix	1.14 ⁹	83 ⁹	AL	76	29	FUKUSHIMA
5	AMIS	0.24	see appendix	1.05	145	AL	200	24	SELLAFIELD

Based on this table such analysis can be made for the models applied to different sites.

Norwegian Fen forest site

Model CROM is highly conservative, as it overestimates both high and low values ($BIAS > 0$ and $MG < 1$). Thus, a relatively high RMSLE is due to a systematic overestimation. NORMALYSA does not behave as a conservative model, as $BIAS < 0$ and $MG > 1$, i.e. predicted values are smaller than measured values on average. Considering these two statistics together with the reported low RMSLE shows that NORMALYSA tends to underestimate systematically, but not very dramatically. The GRANIS model tends to underestimate small values ($MG > 1$), but this is, to some degree, compensated for by the overestimation of high values ($BIAS > 0$). However, a large RMSLE reflects substantial discrepancies in predicting low values.

Thus, CROM is "always on the safe side", but may have the disadvantage of predicting much higher radionuclide contamination (therefore exposure levels) than would actually exist. One goal of TERRITORIES is to reduce the uncertainties of the models. This can be explored further by establishing whether NORMALYSA or GRANIS provided the greater efficacy. If we accept rather small factor of underestimation (recommended applying a correction factor later on), then NORMALYSA might be a good option. GRANIS yields the largest scatter of predicted values relative to measured ones. This behaviour is noteworthy, in the sense that an improved prediction of higher values may be occasional. It is also possible that GRANIS is too sensitive to some input parameter(s), which results in a huge range of predictions. Of course the interpretation for all models is limited bearing in mind that the estimation is based on 5 measured values only (Table 9.2-1).

Belgian NORM site

Although FORESTCROM have a good behaviour in this site, it yields a huge positive bias (overestimation) compared to the other models, but results in rather small overestimation when presented on a logarithmic scale and thus seems to behave less conservatively for low values. ECOFOR behaves almost perfectly both in comparison with the other two models and also based on common criteria, per se, but this is tempered by the observation that the model is calibrated at the same site. We have to keep in mind that ECOFOR is a research model for understanding the processes in a contaminated forest and has not, as yet, been developed as an assessment model. However, there is a potential to develop an easy-to-use assessment tool based on its fundamental concepts. TALAL gives a slightly negative bias and $MG > 1$, i.e. tends to stably underestimate, but not by very much. It may be considered that TALAL is a good model to develop further, but should simulations be needed in a regulatory context there would be a requirement to introduce more conservatism into the model set-up.

Fukushima

For the Fukushima forest site, three models have been used: a simple model using aggregated measured and published transfer factors (FORESTCROM), an already existing forest model: TREE4-simple; and an improved version: TREE4-advanced, where major improvements were achieved in a dynamic process-based forest model. After a calibration with local parameters, the three models exhibit general efficacy, particularly in relation to the behaviour of Cs concentrations in wood. The three models used in this site are not conservative with regards to all the results, whereas a conservative model is expected for many radiation protection purposes.

The FORESTCROM model matches rather well the order of magnitude of observed activity concentrations – on average, the difference (RMSLE) is less than a factor of 2 for needles and wood,

and about a factor of 5 for soil, but it does not perform conservatively in general, as the $BIAS < 0$ and $MG > 0$ show, i.e. modelled values tend to be slightly below measured ones. However, $BIAS > 0$ for wood indicates conservatism for small values. Perhaps FORESTCROM may be considered to provide “uniform” predictions a corollary of which is the underprediction of high values. Fig. 7.2-3 shows that the ranges of measured values in all time periods are huge, thus visual comparison of averages with predictions, resulting in overshooting of concentrations in soil and slight underpredicting in tree organs, remains very indicative, as FORESTCROM does not predict variability.

TREE4-simple model, in contrast, matches better (in a range of two) the concentrations in soil and is less accurate for tree organs and even less so for depuration fluxes (more than a decimal order). However, its predictions are not conservative, as all $BIAS < 0$ and $MG > 0$, but its underprediction of soil concentrations is smaller than that given by FORESTCROM. Fig. 7.2-5 shows that TREE4-simple predicts well higher concentrations in canopy than in bark, whereas the measured values are in the same order in between. Concentrations in wood and bark are clearly underpredicted, which is obviously the reason for the general underprediction for trees (as a summarized measure of the aforementioned components). Underprediction in organic layers occurs at later stages and in 0-20 cm soil layer (also in wood) in the early stage of post-accident evolution.

The TREE4-advanced model managed considerably to reduce the RMSLE for tree organs and depuration fluxes compared to the simpler version. However, the general non-conservative pattern remains. The underestimation is smaller for tree organs and depuration fluxes, but somewhat larger for soil. Fig. 7.2-8 shows that TREE4-advanced performs much more consistently for dynamics in all tree organs than the simple version. It performs rather well also for soil down to 5 cm, but differentiates the deeper layers more than they actually occur. TREE4-advanced seems a good research model to follow the dynamics of radionuclide behaviour in soil and plants but, if there is a need to apply it for regulatory purposes, more conservatism would need to be employed.

Sellafield

The quantitative performance assessment of the ARCTIMAR model shows underestimation of the simulated concentration of Am-241 in filtered water ($BIAS < 0$ and $MG > 1$) and overestimation for the concentration in sediment ($BIAS > 0$ and $MG < 0$). Prediction of the concentration of Cs-137 in water and sediment are overestimated for the high values and underestimated for the low values in comparison with monitoring data ($BIAS > 0$ and $MG > 1$). However, it is necessary to take into account that (i) monitoring data for this region are time vary (taking minimum and maximum values) by 2-3 orders of magnitude, (ii) the main part of ARCTIMAR model simulations are within the range of monitoring data and a significant part of the simulation results for the ARCTIMAR model do not differ more than a factor 2 from monitoring data. Therefore, the RMSLE indicator is relatively low for all estimations.

The quantitative performance assessment of the AMIS model shows a similar pattern of results as ARCTIMAR with the exception of the results for Am-241 for filtered water, where high values are underestimated and low values are slightly overestimated ($BIAS < 0$ and $MG < 1$). For AMIS all statistical indicators show better comparison between simulations and monitoring data than ARCTIMAR. Additionally, values of statistical indicators show relatively good comparison between predictions and monitoring data for the total activity of hot particles for both Am-241 and Cs-137 with (i) slightly overestimation for Am-241 and (ii) underestimation of high values and overestimation of low values for Cs-137.

10 RECOMMENDATIONS

10.1 Selection of a model (prior to performing assessment)

- Select **criteria** that can be used to establish **model adequacy**. Selecting a model that is adequate or **fit-for-purpose** for a given objective is an important step that needs to be performed in advance of any assessment. It seems self-evident that this process would be, among other considerations, dependent on the **context of the assessment** and that this should be informed by stakeholders. A central role in this process should be played by the end user, i.e. the person who will use the outputs of the model and who have all the information related with the application context and purpose of the modelling. The responsibility of the end user, in collaboration with the technical expert in charge of providing simulation results, should extend to selecting the criteria to establish adequacy. These criteria ideally should be specified in terms associated with measurable quantities, such as radionuclide activity concentrations or external exposure (equivalent or absorbed dose), although many times non-measurable quantities are used for selecting the criteria (often effective dose).
- **Uncertainty which the end user is ready to accept** in the assessment is important; if the final result does not have a satisfactory uncertainty outcome, the selection of a more refined model and/or another model may have to be considered.
- The **ability of a model to reproduce measurable data** in the range of application, when data are available for model validation, appears as a prerequisite for its acceptance by the end user. A suitable methodology for comparing correspondence between modelled and empirical datasets has been developed in the present deliverable. Be aware that certain model attributes may not be correlated characteristics. Usually models can be regarded as:
 - More or less conservative models,
 - More or less accurate models, and
 - More or less complex models.

These three characteristics are not correlated. For example, a highly complex model may not necessarily produce more accurate results than a less complex model.

- **Aim towards the simplest practicable model.** In line with ‘Occam’s razor’, it is desirable to develop/adopt a model with the minimum number of assumptions, processes and parameters, which adequately fits with empirical observation. However, in many applications in dose assessments for humans and biota, requiring prediction of radionuclide behavior and fate, accuracy is often achieved by including additional complexity within the models. Conversely, conservatism, in many situations, is sometimes achieved by simplifying the models.
- Be aware that models have **different ranges of application**. Many new models are developed for very specific situations, for instance, to model the dispersion and migration of radionuclides in a specific type of soil with given characteristics (pH, CEC, granulometry, humidity, porosity, chemical composition, microbiota, etc), and transfer to a specific type of vegetation (there are, for example, known important differences in the uptake and translocation of radionuclides in different plant species, e.g. pine, wheat, tomato). This specificity might (although with no guarantee) ensure high accuracy for the given situation, but a slight change in the conditions will give completely different results and may lead to a loss of accuracy. An example of such a case is shown in this report when considering the

coupled tree and soil ‘compartments’, as employed by some of the forest models. The simplest approach used in the report, and widely applied by regulators, is the use of an empirical factor (the concentration ratio) which is element specific, soil dependent and plant species dependent. In the specific case of cesium, a lot of information exists, obtained mainly from data measured in the region affected by the Chernobyl nuclear accident in 1986. However, the application of the same model, using the parameters determined from Chernobyl-contaminated environments, to a different nuclear accident, e.g. Fukushima accident in 2011, provides less than convincing predictions. For this reason, there is a requirement for local parameters to be determined and site-specific models to be developed. Moreover, there are important limitations in the use of more complex models with many chemical parameters, in a new situation, where this information is often not available.

10.2 Use of a model in an assessment

- Consider employing a **tiered or graded approach**. This is especially relevant, with regard to modelling applications within assessments associated with regulation, where the endpoint is to determine whether the risk associated with an exposure is acceptable. As widely used models tend to overestimate the consequences in every situation, the tendency is to develop new models that are able to obtain accurate predictions in particular situations, thereby avoiding an unneeded expense of resources to overprotect populations of humans or biota. This type of approach attributes the level of detail (in e.g. modelling) and resources required for the assessment to the perceived risk associated with the hazard.
- Provide **rigorous evaluation** of the model. This evaluation is obtained by the use of the selected, calibrated and validated model and should address whether the model application meets the objectives of the assessment.
- **Involve stakeholders**. The implementation of a fit-for purpose model is a procedure that should involve all the stakeholders, from the beginning and at all stages of the process, such as those involving the demonstration (including quality assurance) to show that good results have been achieved.

10.3 Developing a model

- Develop a model that is **as simple as possible but able to predict over a wide range of possible conditions**. Ideally, one might use the model to be applicable over a broad number of different compartments in the environment, yielding outputs/results that are adequately (or closely) correlated with empirical measurements. This should be achieved with the highest realism (or accuracy) that is practicable, without losing the possibility of including a degree of conservatism in the case regulators need their use.
- Check that **the typical steps associated with a model development and implementation** procedure are addressed. These steps might include: (1) Model Study Plan, (2) Data and conceptualization, (3) Model set-up, (4) Calibration and validation and (5) Simulation and evaluation.
- **Formulate the problem** and define the assessment context, considering the model might either be **specific** for one given assessment, **or generic** for a range of assessments. It would be typical to ask various questions in relation to the purpose of the assessment(s) and modelling requirements. The availability of relevant data for conducting an assessment might also be considered at this stage. An important task is the analysis of the requirements associated with

the expected accuracy of modelling results. Without defining this at the start of the process, establishing whether a model is fit-for-purpose and whether the appropriate level of complexity has been selected becomes challenging. It is undoubtedly of great utility to define clearly, at the problem formulation stage, the criteria against which the efficacy of the model will be judged and that this list might include information defining an acceptable level of uncertainty in model outputs.

- Efforts should be made to map and **characterise uncertainties** at all stages of the model development where practicable, possibly using some of, but not restricting oneself to, the methods outlined in Appendix 12.2 and those in TERRITORIES D-9.62 (Urso *et al.*, 2019).
- Decide whether the analysis needs to be **deterministic or probabilistic** and, if the latter is required, the model should be developed by using relevant mathematical techniques such as those covered in Urso *et al.* (2019).
- Attempt to document whether all processes are captured. This might involve the development of a conceptual model or the consideration of existing models. One way to achieve this may be through the application of an Interaction Matrix (IM) together with **Features-Event Processes (FEPs) analysis** as described in Section 3.1.
- Attempt to keep the **number of independent model parameters as low as practicable**. Adequately representing mathematically the processes to an adequate level of complexity is a very important challenge. In cases where various models are available for a given radioecological problem the model with the optimum structure should be selected.
- Transform the conceptual model into a **mathematical representation** and **computer coding** in a rigorous manner. That means not only determination of the different mathematical equations for every process, but also appropriate characterization of every parameter included in every equation, in most cases site-dependent to obtain enhanced accuracy.
- Provide appropriate model **calibration and validation**. This might involve obtaining locally determined parameters and input data for the calibration, a comparison with measurements which is part of the validation of the model within the bounds of the applicability and uncertainties.
- Select appropriate **model performance indicators**. In order to test the goodness-of-fit of a given model in a given situation, several metrics exist, and in this report a guidance have been provided, mainly the Root Mean Squared Logarithmic Error (RMSLE) for quantitative performance measurements of a given model. Other metrics may also be applied, regarding the capability of the model to avoid overestimation or underestimation of the data and to preserve data pattern. In addition to these metrics, qualitative indicators are necessary for the purpose of judging the quality of a model. Both qualitative and quantitative indicators can be used to assess model performance. The possibility exists that the model does not meet the requirements set out in the model study plan, for example, the model may not provide the (pre-defined) level of accuracy required or lack coverage of processes considered essential in characterising the system.
- Ensure adequate **quality assurance**. Obviously, an adequate quality assurance is needed in all the steps to establish whether a model is fit for purpose and that the correct level of complexity has been selected.
- **Involve stakeholders**. The development of a fit-for purpose model is a procedure that should involve all the stakeholders, from the beginning and at all stages of the process, such as those

involving (i) the establishment of the desired level of accuracy or conservatism, (ii) the validation of the model to show how accurate the model behaves under different situations and (iii) the demonstration (including quality assurance) that a robust system has been developed.

11 BIBLIOGRAPHY

- American Statistical Association (2016). Statement on statistical significance of p-values [<http://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108#.Vt2XIOaE2MN>].
- Arnold, J. G. M., Daniel N.; Gassman, Philip W.; Abbaspour, Karim C.; White, Michael J.; Srinivasan, Raghavan; Santhi, Chinnasamy; Harmel, Daren; van Griensven, Ann; Van Liew, Michael W.; Kannan, Narayanan; and Jha, Manoj K., . (2012). "SWAT: Model use, calibration, and validation" (2012). . Biological Systems Engineering: Papers and Publications., 406.
- Avila R., Broed R., Pereira A. (2005). ECOLEGO - A toolbox for radioecological risk assessment Proceedings of the International Conference on the Protection from the Effects of Ionizing Radiation, IAEA-CN-109/80. Stockholm: International Atomic Energy Agency. 229 - 232.
- Avila R., Moberg L. (1999). A systematic approach to the migration of ¹³⁷Cs in forest ecosystems using interaction matrices. J. Environ. Radioact., 45 271-282.
- Avila,R., Johansson,E. Bugai, D., Koliabina D. (2018). User's manual for NORMALYSA v.2.1 Description of Program Module Libraries, Mathematical Models and Parameters. Technical Report (for Facilia AB, Sweden) · July 2018; DOI: 10.13140/RG.2.2.16037.40168.
- Barford N. C. (1967). Experimental Measurements: Precision, Error and Truth. Addison-Wesley Publishing Company, Inc.
- Bartos, M.J. and M.R. Palermo (1977). Physical and engineering properties of hazardous industrial wastes and sludges. US EPA Report EPA-600/2-77-139, 77 pp.
- Bennet N. D., Croke B. F. W., Guariso G., *et al.* (2013). Characterizing performance of environmental models. Environmental Modelling & Software 40, 1-20.
- Berg, Ø., Bjørnstad, T., Dahlgren, S., Nøvik, S., Rondeel, W., Totland, A. (2012). Thorium - En framtidssressurs i Oslofjordregionen? Thorium Think Tank rapport til Oslofjordfondet. Rapport nr 2, Regiongeologen, Buskerud Telemark Vestfold fylkeskommuner, Norge, 24 pp.
- BNFL (1976-2016). Discharges and monitoring of the environment in the UK. Environment, Health Safety and Quality, BNFL. 1976-2016, Risley, Warrington.
- Boddy R., Smith G. (2009). Statistical methods in practice: for scientists and technologists. Wiley & Sons. pp. 95–96. ISBN 978-0-470-74664-6.
- Børretzen, P & Salbu, B. (2000). Estimation of apparent rate coefficients for radionuclides interacting with marine sediments from Novaya Zemlya. The Science of the Total Environment, 262, 91-102.
- Brission, N., B. Seguin, and P. Bertuzzi (1992). Agrometeorological soil water balance for crop simulation models. Agricultural and Forest Meteorology. 59: p. 267–287.
- Brooks, R.H. and A.T. Corey (1966). Properties of porous media affecting fluid flow. Soil Science Society of America Journal. 39(2): p. 61– 88.
- Brown, J.E., Alfonso, B., Avila, R., Beresford, N.A., Coppleson, D., Hosseini, A. (2016). A new version of the ERICA tool to facilitate impact assessments of radioactivity on wild plants and animals. J. Environ. Radioact. 153, 141–148.
- Burnham K. P., Anderson D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd edition. Springer-Verlag. ISBN 0-387-95364-7.
- Calder, M., Craig, C., Culley, D., de Cani, R., Donnelly, C. A., Douglas, R., . . . Wilson, A. (2018). Computational modelling for decision-making: where, why, what, who and how. R Soc Open Sci, 5(6), 172096. doi: 10.1098/rsos.172096

- Calmon P, Thiry Y., Zibold G., Rantavaara A., Fesenko S. (2009). Transfer parameter values in temperate forest ecosystems: a review. *Journal of Environmental Radioactivity* 100: 757-766.
- Calmon P. (2014). TREE4: Transfer of Radionuclides and External Exposure in FORrest systems. Technical Report IRSN/SESURE/2014-001 (in French)
- Calmon, P., M.A. Gonze and C. Murlon (2015). Modeling the early-phase redistribution of radiocaesium fallouts in a evergreen coniferous forest after Chernobyl and Fukushima accidents. *Science of the Total Environment* 529: p. 30-39.
- Campbell, G.S. (1974). Simple method for determining unsaturated conductivity from moisture retention data. *Soil Science*, 1974. 117: p. 311-314.
- Casadesus, J., T. Sauras-Yera, and V.R. Vallejo (2008). Predicting soil-to-plant transfer of radionuclides with a mechanistic model (BioRUR). *Journal of Environmental Radioactivity*, 2008. 99(5): p. 864-871.
- Chai T., Draxler R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* 7, 1247-1250.
- Chakravarti I. M., Laha R. G., Roy J. (1967). *Handbook of Methods of Applied Statistics*, Volume I. John Wiley. pp. 392-394.
- Chang J. C., Hanna S. R. (2005). *Technical Descriptions and User's Guide for the BOOT Statistical Model Evaluation Software Package, Version 2.0*. George Mason University, Homeland Security Institute & Harvard School of Public Health, 62 p. [http://www.harmo.org/kit/Download/BOOT_UG.pdf].
- Chapman, N.A.; Andersson, J.; Robinson, P.; Skagius, K.; Wene, C-O.; Wiborgh, M.; and Wingefors, S. (1995). *Systems Analysis, Scenario Construction and Consequence Analysis Definition for SITE-94*. SKI Report 95:26. Stockholm, Sweden: Swedish Nuclear Power Inspectorate
- Deckmyn, G., *et al.* (2008). ANAFORE: a stand-scale process-based forest model that includes wood tissue development and labile carbon storage in trees. *Ecological Modelling*, 215: p. 345-368.
- Deckmyn, G., *et al.* (2011). Simulating C cycles in forest soils: Including the active role of micro-organisms in the ANAFORE forest model. *Ecological Modelling*, 222: p. 1972-1985.
- Diener, A., Hartmann, P., Urso, L., i Batlle, J. V., Gonze, M., Calmon, P., & Steiner, M. (2017). Approaches to modelling radioactive contaminations in forests—Overview and guidance. *Journal of Environmental Radioactivity*, 178, 203-211.
- Dowdall, M., Brown, J.E., Hosseini, A. and Mora Cañadas J.C. (2012). Impact of legacy enhanced natural radioactivity on human and natural environments. In : *Radionuclides: Sources, Properties and Hazards*. Nova Science Publishers, Inc. pp. 171-204.
- Dvorzhak A., Puras C., Montero M. and Mora J.C. (2012). Spanish Experience on Modeling of Environmental Radioactive Contamination due to Fukushima Daiichi NPP Accident Using JRODOS. *Environmental Science & Technology* 46(21), p. 11887-11895.
- EC, European Commission (2003). *Radiation protection 132. MARINA II, Update of the MARINA Project on the radiological exposure of the European Community from radioactivity in North European marine waters*. Luxembourg: Office for Official Publication of the European Communities.
- Efron B., Tibishirani R. (1993). *An Introduction to the Bootstrap*. Chapman & Hall.
- Ehrhardt J. and Weiss A. (2000). *RODOS: Decision support for off-site nuclear emergency management in Europe*. - EUR19144EN, Luxemburg, European Community.
- EPA (1993). *External exposure to radionuclides in air water and soil* (by Eckerman K. F., Ryman J.C.), Federal Guidance Report No. 12, Office of Radiation and Indoor Air U.S. Environmental Protection Agency (EPA), Washington, DC.
- Gelfand A. E., Ghosh S. K. (1988). Model choice: a minimum posterior predictive loss approach. *Biometrika* 85 (1), 1-11.
- Ghelli A., Ebert E. (2008). Special issue on forecast verification. *Meteorological Applications* 15 (1), 1350-4827 [<https://doi.org/10.1002/met.69>].
- Gielen, S., *et al.*, Concentrations and distributions of Al, Ca, Cl, K, Mg and Mn in a Scots pine forest in Belgium. *Ecological modelling*, 2016. 324: p. 1-10.

- Goktas, R.K. and M.M. Aral (2011). Integrated Dynamic Modeling of Contaminant Fate and Transport within a Soil–Plant System. *Vadose Zone Journal*. 10: p. 1130-1150.
- 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.
- Gonze M.-A., Mourlon C., Calmon P., Manach E., Debayle C., Baccou J. (2016). Modeling the Dynamics of Ambient Dose Rates Induced by Radiocaesium in the Fukushima Terrestrial Environment. *Journal of Environmental Radioactivity* 161: 22-34
- Goussev, N.G., E. Kovalev, and A. Foderaro (1968). Derivation of formulae, in *Engineering Compendium on Radiation Shielding*, R.G. Jaeger, Editor. Springer and Verlag: New York.
- Goussev, N.G. (1968). Relationship between absorbed dose rate (in air) and particle flux density or energy flux density (X- and gamma rays), in *Engineering Compendium on Radiation Shielding* R.G. Jaeger, Editor. Springer and Verlag: New York.
- Granier, A., *et al.* (1999). A lumped water balance model to evaluate duration and intensity of drought constraints in forest stands. *Ecological Modelling*. 116(2-3): p. 269-283.
- Guillevic J., *et al.* (2018). CONCERT TERRITORIES deliverable report D9.65. https://territories.eu/assets/files/publications/D9-65_TERRITORIES_Decision%20processes_pathways_final_30012018_approved.pdf
- Harmel R.D *et al.* (2006). Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Transactions of the ASABE* 49:689-701.
- Hastie T., Tibshirani R., Friedman J. (2009). *The Elements of Statistical Learning*. Springer Series in Statistics, 2nd edition.
- HM TREASURY (2013). Review of quality assurance of government models. <http://www.gov.uk/government/publications/review-of-quality-assurance-of-government-models>.
- IAEA (2001). *Generic Models for Use in Assessing the Impact of Discharges of Radioactive Substances to the Environment*. IAEA Safety Reports Series 19, 216. STI/PUB/1102.
- IAEA (2004). *Sediment distribution coefficients and concentration factors for biota in the marine environment*. IAEA Technical reports series no. 422. Vienna: International Atomic Energy Agency, IAEA.
- IAEA (2009)., *Quantification of radionuclide transfer in terrestrial and freshwater environments for radiological assessments*, IAEA-TECDOC-1616, Vienna, 616 pp. 2009.
- IAEA (2010). *Handbook of Parameter Values for the Prediction of Radionuclide Transfer in Terrestrial and Freshwater Environments*. Technical Report Series No. 472. The International Atomic Energy Agency, Vienna.
- IAEA (2014), *Handbook of Parameter Values for the Prediction of Radionuclide Transfer to Wildlife*. Technical Reports Series No. 479. International Atomic Energy Agency. Vienna, 2014.
- IAEA (2018). *Regulatory control of radioactive discharges to the environment / International Atomic Energy Agency*. Description: Vienna : International Atomic Energy Agency, 2018. | Series: IAEA safety standards series, ISSN 1020–525X ; no. GSG-9 |
- ICRP (1996). *Conversion Coefficients for use in Radiological Protection against External Radiation*. ICRP Publication 74. Ann. ICRP 26 (3-4). 1996.
- ICRP (2006). *Assessing Dose of the Representative Person for the Purpose of Radiation Protection of the Public (101A) and The Optimisation of Radiological Protection: Broadening the process*. ICRP Publication 101 - Ann. ICRP 36 (3), pp 112.
- ICRP (2007). *The 2007 Recommendations of the International Commission on Radiological Protection*. ICRP publication 103. *Annals of the ICRP*, 37(2-4).
- IFE (2006). *Radiologisk kartlegging av området rundt tidligere Sørve gruver*. Institutt for energiteknikk (IFE); IFE Report, IFE/KR/F - 2006/174 (In Norwegian).
- Imamura, N., M. Komatsu, S. Ohashi, S. Hashimoto, T. Kajimoto, S. Kaneko and T. Takano (2017). "Temporal changes in the radiocesium distribution in forests over the five years after the Fukushima Daiichi Nuclear Power Plant accident." *Scientific Reports* 7(1): 8179.

- Iosjpe M. (2006). Environmental Modelling: Modified Approach for Compartmental Models. In: Radionuclides in the Environment. Edd. P.P. Povinec, J.A. Sanchez-Cabeza. Radioactivity in the Environment, vol. 8, Series Editor: M.S. Baxter, 2006, 463-476.
- Iosjpe M., Brown J. & Strand P. (2002). Modified Approach for Box Modelling of Radiological Consequences from Releases into Marine Environment, Journal of Environmental Radioactivity, Vol. 60, No 1-2, 91-103.
- Iosjpe M., M. Karcher, J. Gwynn, I. Harms, R. Gerdes and F. Kauker (2009). Improvement of the dose assessment tools on the basis of dispersion of the 99Tc in the Nordic Seas and the Arctic Ocean. Radioprotection, vol. 44, no 5, 531-536.
- Iosjpe, M. (2011). A sensitivity analysis of the parameters controlling water-sediment interactions in the coastal zone: Consequences to man and environment. Journal of Marine Systems, 2011. 88: p. 82-89.
- Iosjpe, M. (2006) Environmental Modelling: Modified Approach for Compartmental Models. In: Radionuclides in the Environment. Edd. P.P. Povinec, J.A. Sanchez-Cabeza. Radioactivity in the Environment, vol. 8, Series Editor: M.S. Baxter, 2006, 463-476. 2006.
- Iosjpe, M., *et al.* (2009). Improvement of the dose assessment tools on the basis of dispersion of the 99Tc in the Nordic Seas and the Arctic Ocean. Radioprotection. 44(5): p. 531-536.
- Iosjpe, M. (2011). Evaluation of the environmental parameters controlling the vulnerability of the coastal marine regions. Radioprotection. 46(6): p. S289-S293.
- Iosjpe, M., J. Brown, and P. Strand (2002). Modified Approach for Box Modelling of Radiological Consequences from Releases into Marine Environment. Journal of Environmental Radioactivity. 60(1-2): p. 91-103.
- Jones K. *et al.* (2019). CONCERT-TERRITORIES deliverable report D9.63 https://territories.eu/assets/files/publications/D9.63-Guidance-about-exposure-scenarios_approved04072019.pdf
- Jørgensen S.E. (1994). Fundamentals of ecological modelling. Developments in Environmental Modelling, 19, Elsevier, Amsterdam-London-New York-Tokyo, 1994.
- Karcher M.J. and Harms I.H. (2000). Estimation of water and ice fluxes in the Arctic for an improved box structure of the NRPA box model. In: Iosjpe M. (Ed.): Transport and fate of contaminants in the Northern Seas, NRPA, 2000.
- Kendy, E., *et al.* (2003). A soil-water balance approach to quantify groundwater recharge from irrigated cropland in the North China Plain. Hydrological Processes. 17: p. 2011-2031.
- Kirchner, G., & Steiner, M. (2008). Uncertainties in radioecological assessment models--their nature and approaches to reduce them. Appl Radiat Isot, 66(11), 1750-1753. doi: 10.1016/j.apradiso.2007.09.021.
- Kohavi R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence Vol. 2, pp. 1137-1143.
- Kowe, R., *et al.* (2007). GRANIS: A Model for the Assessment of External Photon Irradiation from Contaminated Media of Infinite Lateral Extent. Chilton.
- Krumbein, M.B. and E. Aberdeen (1937). The sediments of Brataria Bay. Journal of sedimentary Petrology. 7: p. 3-17.
- Ku H.H. (1966). Notes on the use of propagation of error formulas. Journal of research of the national bureau of standards. 70C: 263-273.
- Kuhnert M., Voinov A., Seppelt R. (2006). Comparing raster map comparison algorithms for spatial modelling and analysis. Photogrammetric Engineering and Remote Sensing 71 (8), 685-695.
- Lamb, H. (1994). Hydrodynamics (6th ed.). Cambridge University Press.
- Larsson, C.M., (2008). An overview of the ERICA Integrated Approach to the assessment and management of environmental risks from ionising contaminants. J. Environ. Radioact. 99, 1364-1370.

- Lee & Kim (2017). Post-closure safety assessment of near surface disposal facilities for disused sealed radioactive sources. *Nuclear Engineering and Design*, 313, 425–436.
- Legates R. D., McCabe G. (1999) Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35, 233-241 [<https://doi.org/10.1029/1998WR900018>].
- Liu, C., X. Zhang, and Y. Zhang (2002). Determination of daily evaporation and evapotranspiration of winter wheat and maize by large-scale weighing lysimeter and micro-lysimeter. *Agricultural and Forest Meteorology*. 111: p. 109-120.
- MacKenzie J. and Nicholson S. (1987). COLDOS – A computer code for the estimation of collective doses from radioactive releases to the sea. SRD R 389, UK Atomic Energy Authority.
- Maître M. *et al.* (2019). CONCERT-TERRITORIES Deliverable report D9.66. https://territories.eu/assets/files/publications/D9.66_Stakeholders-panels-results-France_approved04072019.pdf
- Matott, L. S., Babendreier, J. E., & Purucker, S. T. (2009). Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resources Research*, 45(6). doi: 10.1029/2008wr007301
- Mitchell P.I., Condren O.M. Leon Vintro L. & McMahon C.A. (1999). Trends in plutonium, americium and radiocaesium accumulation and long-term bioavailability in the western Irish Sea mud basin. *Journal of Environmental Radioactivity*, Vol. 44, 223-251.
- Monteith, J. and M. Unsworth (2007). *Principles of Environmental Physics*, Third Ed. Academic press, 418 pp.
- Mora, J. C. *et al.* (2018). Application of existing models to the TL. Internal Report (MS1.3, task 1.2) to the TERRITORIES Project.
- Mora, J.C *et al.* (2015). Generic Model for Combined Tier-1. Assessments for Humans and Wildlife; STAR (Contract Number: Fission-2010-3.5.1-269672) DELIVERABLE D-Nº 3.1; Publications Office of the European Union: Rue Mercier, Luxembourg, 2015.
- Mora, J. C. *et al.* (2018). CONCERT-TERRITORIES deliverable report D9.60.
- Moriasi D. N., Arnold J. G., Van Liew M. W., *et al.* (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers* 50 (3), 885-900.
- Mrdakovic Popic J., Salbu B., Strand T., Skipperud, L. (2011). Assessment of radionuclides and metal contamination in thorium rich area in Norway. *J.Environ.Monit.* 13, 1730-1738.
- Mrdakovic Popic J., Bhatt C.R., Salbu B., L.Skipperud (2012). Outdoor 220Rn, 222Rn and terrestrial gamma radiation levels: investigation study in the thorium rich Fen Complex, Norway. *J.Environ.Monit.* 14, 193-201.
- Mrdakovic Popic J. (2014). Doctoral thesis: Environmental impact of radionuclide and trace elements in the thorium rich Fen area in Norway. Norwegian University Of Life Sciences, Norway.
- Nash J. E., Sutcliffe J. V. (1970). River flow forecasting through conceptual models, Part I - A discussion of principles. *Journal of Hydrology* 10, 282–290.
- NGI-UMB (2010). Kartlegging av omfang og kostnader ved eventuell senere opprydning av radioaktivt materiale ved Søve gruver; Norwegian Geotechnical Institute (NGI) and Norwegian University of Life Sciences, Rapport 1927-00-14-R, pp. 157 (In Norwegian).
- Nielsen S.P., Iosjpe M. & Strand P. (1997). Collective doses to man from dumping of radioactive waste in the Arctic Seas. *The Science of the Total Environment*, 202, 135-146.
- Nielsen, S.P. and M. Keith-Roach (2004). Update of the MARINA Project on the radiological exposure of the European Community from radioactivity in North European marine waters. Annex E: Critical group exposure. European Commission report, 30 pp.
- Nielsen, S.P., M. Iosjpe, and P. Strand (1997). Collective doses to man from dumping of radioactive waste in the Arctic Seas. *Science of the Total Environment*. 202: p. 135-146.
- Official Journal of the European Union (2014) European Council Directive 2013/59/Euratom on basic safety standards for protection against the dangers arising from exposure to ionising radiation

- and repealing directives 89/618/Euratom, 90/641/Euratom, 96/29/Euratom, 97/43 Euratom and 2003/122/Euratom. OJ of the EU. L13;57:1-73 Available at <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2014:013:0001:0073:EN:PDF>. Accessed 11 Jul 2019
- Ohashi S., Katsushi Kuroda, Tsutomu Takano, Youki Suzuki, Takeshi Fujiwara, Hisashi Abe, Akira Kagawa, Masaki Sugiyama, Yoshitaka Kubojima, Chunhua Zhang, Koichi Yamamoto (2017). Temporal trends in ¹³⁷Cs concentrations in the bark, sapwood, heartwood, and whole wood of four tree species in Japanese forests from 2011 to 2016. *Journal of Environmental Radioactivity* 178-179 335-342. And supplementary material.
- Periáñez, R. (2003). Redissolution and long-term transport of radionuclides released from a contaminated sediment: a numerical modelling study. *Estuarine, Coastal and Shelf Science* 56: p. 5-14.
- Periáñez, R. (2003) Kinetic modelling of the dispersion of plutonium in the eastern Irish Sea: two approaches. *Journal of Marine Systems*,. 38(3-4): p. 259-275.
- Periáñez R. (2005). Modelling the dispersion of radionuclides in the marine environment. Springer-Verlag, Heidelberg.
- Periáñez, R., *et al.* (2018). The marine kd and water/sediment interaction problem. *Journal of Environmental Radioactivity* 192: p. 635-647.
- Ramsey C. A., Hewitt A. D. (2005). A methodology for assessing sample representativeness. *Environmental Forensics* 6, 71-75.
- Rantavaara, A., *et al.* (2001). Forest food chain and dose model (FDMF) for RODOS. Model description. STUK report – A178. Dec 2001, 65 pp.
- Refsgaard, J. C., Henriksen, Hans Jørgen, Harrar, William G., Scholten, Huub, Kassahun, Ayalew. (2005). Quality assurance in model based water management – review of existing practice and outline of new approaches. *Environmental Modelling & Software*, 20(10), 1201-1215. doi: 10.1016/j.envsoft.2004.07.006
- Refsgaard, J. C., van der Sluijs, Jeroen P., Højberg, Anker Lajer, Vanrolleghem, Peter A. (2007). Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software*, 22(11), 1543-1556. doi: 10.1016/j.envsoft.2007.02.004.
- Reich S., Cotter C. (2015). Probabilistic Forecasting and Bayesian Data Assimilation. Cambridge University Press.
- Richards, L.A. (1931). Capillary conduction of liquids through porous mediums. *Physics*. 1: p. 318–333.
- RIFE (1995-2014). RIFE reports: liquid and gaseous discharges from nuclear sites in England and Wales from 1995 to 2014.
- Roberts, J. (1976). A study of root distribution and growth in a *Pinus sylvestris* L. (Scots pine) plantation in East Anglia. *Plant and Soil*. 44: p. 607-621.
- RODOS (1999). Model description of the terrestrial food chain and dose module FDMT in RODOS PV 4.0. RODOS(WG3)-TN(99)17.
- Ross S. M. (2017). *Introductory Statistics*, 4th edition. Academic Press. ISBN 9780128043172.
- Rühm W., Steiner M., Kammerer L., Hiersche L., Wirth E. (1998). Estimating future radiocaesium contamination of fungi on the basis of behaviour patterns derived from past instances of contamination. *Journal of Environmental Radioactivity* 39 (2), 129-147.
- Santhi C., Arnold J. G., Williams J. R., *et al.* (2001). Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of American Water Resources Association* 37 (5), 1169–1188.
- Savage D, Maul PR, Benbow S, Walke RC (2004). A generic FEP database for the assessment of long-term performance and safety of the geologic storage of CO₂. Quintessa Report QRS-1060A-1; 73pp (http://www.quintessa-online.com/FED_DB_report.pdf)
- Schwarz Gideon E. (1978). Estimating the dimension of a model. *Annals of Statistics* 6 (2), 461–464 [<https://doi.org/10.1214/aos/1176344136>].

- Sellafield Ltd (2011). Particles in the Environment. Annual Report for 2010/2011 and Forward Programme, Whitehaven: Nuclear Decommission Authority.
- Sellafield Ltd (2014). Particles in the Environment. Annual Report for 2013/2014 and Forward Programme, Whitehaven: Nuclear Decommission Authority.
- Sellafield Ltd (2015). Particles in the Environment. Annual Report for 2014/2015 and Forward Programme, Whitehaven: Nuclear Decommission Authority.
- Sellafield Ltd (2016). Particles in the Environment. Annual Report for 2015/2016 and Forward Programme, Whitehaven: Nuclear Decommission Authority.
- Sellafield Ltd (2017). Particles in the Environment. Annual Report for 2016/2017 and Forward Programme, Whitehaven: Nuclear Decommission Authority.
- Sellafield Ltd (2018). Particles in the Environment. Annual Report for 2017 and Forward Programme, Whitehaven: Nuclear Decommission Authority.
- Shaw G., A. Venter, R. Avila, R. Bergman, A. Bulgakov, P. Calmon, S. Fesenko, M. Frissel, F. Goor, A. Konoplev, I. Linkov, S. Mamikhin, L. Moberg, A. Orlov, A. Rantavaara, S. Spiridonov, Y. Thiry (2005). Radionuclide migration in forest ecosystems – results of a model validation study. *Journal of Environmental Radioactivity* 84: 285-296.
- Singh, J., H. V. Knapp, and M. Demissie (2004). Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004-08. Champaign, Ill.: Illinois State Water Survey. Available at: www.sws.uiuc.edu/pubdoc/CR/ISWSCR2004-08.pdf. Accessed 8 September 2005.
- Smith J. *et al.* (2019). D 9.59 - Radiological state database of sites. CONCERT- TERRITORIES Deliverable report.
- Spiegelhalter D. J., Best N. G., Carlin B. P., Linde A. V. D. (2002). Bayesian measures of model complexity and fit. *Journal of Royal Statistical Society Ser. B Stat. Methodol.* 64 (4), 583-639.
- Sweeck L., Bradshaw C., Mora J.C., Stark K., Beaugelin K., Beresford N., Boyer P., Fevrier L., Hurtevent P., Outola L., Robles B., Steiner M., Thørring H., Turtiainen T., Vetikko V., Vives i Batlle J. (2014). STAR internal report MS 35: Report analysing both FEP and IM analysis with respect to a conceptual model for integrated risk assessment. Minutes from 2-3 of May videoconference meeting. STAR project Grant agreement no: 269672. May 2014.
- Sy M. M., Ancelet S., Henner P., Hurtevent P., Simon-Cornu M. (2015). Foliar interception of radionuclides in dry conditions: a meta-analysis using a Bayesian modeling approach. *Journal of Environmental Radioactivity* 147, 63-75.
- Thorium Committee (2008). Report: Thorium as an energy source - opportunities for Norway. Thorium Report, ISBN 978-82-7017-692-2 (printed) ISBN 978-82-7017-693-9, 150 pp.
- UNSCEAR (2000). Sources and Effects of Ionizing Radiation. Report to the General Assembly, with scientific annexes. United Nations, New-York. 2000.
- Urso, L. *et al.* (2019). D 9.62 – Guidance on uncertainty analysis for radioecological models. CONCERT- TERRITORIES Deliverable report.
- USDoE (2002). A Graded Approach for Evaluating Radiation Doses to Aquatic and Terrestrial Biota. U.S. Department of Energy, Washington D.C., U.S.A. Technical Standard DOE-STD-1153-2002.
- Van den Hoof, C. and Y. Thiry, Modelling of the natural chlorine cycling in a coniferous stand: implications for chlorine-36 behaviour in a contaminated forest environment. *Journal of Environmental Radioactivity*, 2012. 107: p. 56–67.
- Van Genuchten, M.T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 1980. 44(5): p. 892–898.
- Van Liew M. W., Arnold J. G., Garbrecht J. D. (2003). Hydrologic simulation on agricultural watersheds: Choosing between two models. *Transactions of the American Society of Agricultural Engineers* 46 (6), 1539-1551.
- Vandenhove, H., *et al.* (2009). Proposal for new best estimates of the soil-to-plant transfer factor of U, Th, Ra, Pb and Po. *Journal of Environmental Radioactivity*. 100: p. 721-732.

- Vanhoudt, N. (2015). Sludge heap 'Kepkensberg' from Belgian phosphate industry. European Observatories for Radioecological Research – Template for the Description of Candidate Sites. EC STAR Project Report, 20 pp.
- Velasco H. R., J. J. Ayub, M. Belli, U. (2006). Sansone Interaction matrices as a first step toward a general model of radionuclide cycling: Application to the ¹³⁷Cs behavior in a grassland ecosystem Journal of Radioanalytical and Nuclear Chemistry, Vol. 268, No.3 (2006) 503–509
- Vincke, C. and Y. Thiry (2008). Water table is a relevant source for water uptake by a Scots pine (*Pinus sylvestris* L.) stand: Evidences from continuous evapotranspiration and water table monitoring. Agricultural and Forest Meteorology, 2008. 148(10): p. 1419-1432.
- Vincke, C., *et al.* (2005). Evapotranspiration of a declining *Quercus robur* (L.) stand from 1999 to 2001. I. Trees and forest floor daily transpiration. Annals of Forest Science. 62(6): p. 503-512.
- Vincke, C. (2006). Technical report of the post-doc programme "Contribution to long term biosphere impact modelling in a geological disposal scenario through measurements of soil-tree-atmosphere water transfer and related ecophysiological parameters". Belgian Nuclear Research Centre Report SCK•CEN-I-90, Mol, Belgium, 63 pp. <http://hdl.handle.net/2078.1/95302>.
- Vives i Batlle, J., H. Vandenhove, and S. Gielen (2014). Modelling water and ³⁶Cl cycling in a Belgian pine forest. In: Proc. ICRER 2014 - 3rd International Conference on Radioecology & Environmental Radioactivity, 7-12 September 2014, Barcelona, Catalonia. Electronic proceedings Extended Abstract. Available from: <http://radioactivity2014.pacifico-meetings.com/>.
- Wilk M. B., Gnanadesikan R. (1968). Probability plotting methods for the analysis for the analysis of data. Biometrika 55 (1), 1–17 [<https://doi.org/10.1093/biomet/55.1.1>].
- Willmott C. J. (1981). On the Validation of Models. Physical Geography 2, 184-194.
- Wirth E. (1982). The applicability of the C-14 specific activity model. Health Physics 43, 919-922.
- Xiao, C.W. and R. Ceulemans (2004). Allometric relationships for below- and aboveground biomass of young Scots pines. Forest Ecology and Management. 203: p. 177–186.

12 APPENDICES

12.1 Performance assessment analysis. Old vs New model comparison.

From one hand widely used and from the other more advanced models have been applied to three forest and one marine site and we have reported in this milestone the results of applying them. The characteristics and definitions of those sites are described above.

A statistical analysis of the results has been carried out for each model using the guidance provided in TERRITORIES Milestone MS1.7. The root mean squared logarithmic error (RMSLE), BIAS metric (BIAS) and geometric mean bias (MG) (see correspondent formulas 3.17, 3.7, 3.8 in the Section 3) were used to provide a measure of model performance (see following).

12.1.1 NORWEGIAN FEN FOREST SITE

12.1.1.1 Widely applied models

12.1.1.1.1 Performance assessment analysis of CROM model applied to FEN forest site (Norway)

Table 12.1-1: Quantitative performance assessment of **CROM** applied to the Fen site

Quantitative performance table for CROM applied to the Fen site (Norway)	
Indicators	Values
Root mean squared logarithmic error (RMSLE)	0.59
BIAS	2.46
Geometric mean bias (MG)	0.69
Cross-correlation function	There are no time series in the database to perform the test

Table 12.1-2: Qualitative performance assessment of CROM applied to the Fen site

Qualitative performance table for CROM applied to the Fen site (Norway)	
Indicators	Values
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> • CROM code implements IAEA SRS 19 (IAEA, 2001) generic models for dispersion and transfer of radionuclides. The models implement more than 100 parameters, most of them radionuclide dependent (at this moment 162 radionuclides implemented). • For the specific case of Fen, only measurements of ambient dose equivalent were compared with the models implemented in CROM. Dose Conversion Factors (DCFs), for this specific site, are implemented in the code for converting radioactivity deposition on the soils (Bq/m²/d) into ambient dose equivalent (Sv/y). These factors are calculated using predefined density, grain size and chemical composition of the soil, and hypothesis on the period for the decay chain ingrowth (see FGR12). Surface deposit is considered. • Additional hypothesis were assumed for the factor to convert the measurements from Fresh Weight (FW) to Dry Weight (DW), for the transformation from Kerma to ambient dose equivalent and for the correction of cosmic rays. • With these hypotheses, 7 DCFs (based in some fixed parameters) were used for the measured radionuclides.
Number of processes considered separately	<ul style="list-style-type: none"> • All the transport processes for particles considered in MCNP (model used for the derivation of the DCFs) and the decay processes used in the bateman equations to calculate the accumulation in the period considered (30 years).
Application of the model to past set of data	<ul style="list-style-type: none"> • The models in CROM have been widely applied to many data sets in the past. For instance the atmospheric Gaussian plume model is a the generic model implemented, which has been validated against many data sets for more than 50 years.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> • Many of the models in CROM can be applied to other chemical stable elements. • For the specific application in Fen site, the dose conversion factors are only derived and can be only applied to radioactive elements.

12.1.1.1.2 Performance assessment analysis of NORMALYSA model applied to FEN forest site (Norway)

The comparison in Table 12.1-3 is for gamma air kerma rate as predicted by the model NORMALYSA and measured in situ at Fen.

Table 12.1-3: Quantitative performance assessment of *NORMALYSA* applied to the Fen site

Quantitative performance table for <i>NORMALYSA</i> applied to the Fen site (Norway)	
Indicators	Values
Root mean squared logarithmic error (RMSLE)	0.23
BIAS	-0.23
Geometric mean bias (MG)	1.37
Cross-correlation function	There are no time series in the database to perform the test

Table 12.1-4: Qualitative performance assessment of *NORMALYSA* applied to the Fen site

Qualitative performance table for <i>NORMALYSA</i> applied to the Fen site (Norway)	
Indicators	Values
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> For the specific case of Fen, only measurements of ambient dose equivalent were compared with the models implemented in <i>NORMALYSA</i>. Dose coefficients for effective doses from external irradiation from surface deposition are based on [EPA, 1993]. Dose coefficients are calculated using predefined density; grain size and chemical composition of the soil, and hypothesis on the period for the decay chain ingrowth (see EPA, 1993). For a single point in time (no radionuclide migration and redistribution processes modelled), the number of parameters is limited to basic soil properties – including waste and cover: bulk density, porosity, moisture content, depth. Dose coefficients can be altered although this would require justification. When comparing observed and predicted data in a normalized way, information is required with respect to the cosmic component of radiation and an ambient dose to air kerma conversion factor. 9 Dose coefficients (based on fixed parameters) were used for the measured radionuclides. There is no time dependency in this simulation as instantaneous, at time = 0, observed and predicted values are compared.
Number of processes considered separately	<ul style="list-style-type: none"> All the radiation transport processes for gamma photons considered in MCNP (model used for the derivation of the Dose coefficients). The model can account for ingrowth and decay of radionuclides within the U-238 and Th-232 decay chains but for our purposes secular equilibrium was assumed for daughters where empirical data were not available. The process of shielding by a cover layer can be included but this was set to zero for the sake of conservatism.
Application of the model to past set of data	<ul style="list-style-type: none"> <i>NORMALYSA</i> actually comprises of a suite of models, many of which have been applied to and tested against empirical datasets over many decades. As noted above, the external exposure dose coefficients are from EPA (1993) and have been widely applied in many studies world-wide since publication (e.g. IAEA, 2001).

Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> • NORMALYSA has focus on natural decay series radionuclides and as such the application to model the behavior and fate of stable isotopes and chemical analogues would have limited value (not least because most of the radionuclides considered, with the exception of Pb, Bi and Tl, don't have stable analogues). • For the specific application at the Fen site, the dose coefficients are selected as defaults and can be only applied to radioactive elements.
---	--

12.1.1.2 More advanced models

12.1.1.2.1 Performance assessment analysis of GRANIS model applied to FEN forest site (Norway)

Table 12.1-5: Quantitative performance assessment of GRANIS applied to the Fen site

Quantitative performance table for GRANIS applied to the Fen site (Norway)	
Indicators	Values
Root mean squared logarithmic error (RMSLE)	0.77
BIAS	0.47
Geometric mean bias (MG)	2.65
Cross-correlation function	There are no time series in the database to perform the test

Table 12.1-6: Qualitative performance assessment of GRANIS applied to the Fen site

Qualitative performance table for GRANIS applied to the Fen site (Norway)	
Indicators	Values
Number of parameters involved and their time-independence	<p>Section 5.2.1 describes the GRANIS model. Parameters include:</p> <ul style="list-style-type: none"> • Height at which dose is calculated • Elemental composition of contaminated and shielding material • Density and thickness of materials • Build-up factors which depend primarily on the depth and composition (specifically atomic number of elements) of materials considered • Fluence to dose conversion library • Radionuclides in decay chain • Photon energy and intensities • Attenuation lengths for photons • Measured activity concentration in layers <p>Can be linked to a soil model to predict the migration of radionuclides in soil and resulting doses as a function of time but this time dependence is not considered here.</p>

Number of processes considered separately	<ul style="list-style-type: none"> • Attenuation of photons of different energies in different media • Scattering using build-up factors • Calculation of particle flux at distance of interest • Calculation of absorbed dose in air
Application of the model to past set of data	<ul style="list-style-type: none"> • The report HPA-RPD-032* describes inter-comparisons of the results of GRANIS v3 with standard theoretical methods, published papers and other radiation transport codes. GRANIS v3 was found to agree well with other results for photon energies in the range 0.1 to 4 MeV - photon energies of importance for radionuclides found in the environment.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> • Not applicable.

*https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/340113/HpaRpd032.pdf

12.1.2 BELGIAN NORM SITE

12.1.2.1 Widely applied models

12.1.2.1.1 Performance assessment analysis of simple forest model FORESTCROM applied to the Belgian NORM site

Table 12.1-7: Quantitative performance assessment of **FORESTCROM** applied to the Belgian NORM site

Quantitative performance table for simple forest model FORESTCROM applied to the Belgian NORM site	
Indicators	Values
Root mean squared logarithmic error (RMSLE)	4.57 × 10 ⁻²
BIAS	11.4
Geometric mean bias (MG)	0.90
Cross-correlation function	There are no time series in the database

Table 12.1-8: Qualitative performance assessment of *FORESTCROM* applied to the Fen site

Qualitative performance table for simple forest model FORESTCROM applied to the Belgian NORM site	
Indicators	Values
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> The mathematical model was developed for TERRITORIES project to show how a simple model for a forest could be derived. In general, the model involves 12 parameters for each radionuclide considered in a site. For the specific case of the Belgian NORM site, no deposition on the external surface of the trees, or losses of the radionuclides in the soil were considered. Therefore the number of parameters was reduced to 5 for each radionuclide. Although the model is dynamic, all the parameters are independent of the time.
Number of processes considered separately	<ul style="list-style-type: none"> Three main processes are considered in the model, as simplified for this site, include the root uptake (based in locally determined parameters), the radioactive decay and an ecological decay period (which include processes as the litter-fall).
Application of the model to past set of data	<ul style="list-style-type: none"> The model has been developed for this project, and no back-testing has been applied at this moment.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> The model can be easily adjusted to stable elements.

12.1.2.2 More advanced models

12.1.2.2.1 Performance assessment analysis of ECOFOR model applied to Belgian NORM site

Table 12.1-9: Quantitative performance assessment of *ECOFOR* applied to the Belgian NORM site

Quantitative performance table for ECOFOR (applied to the Belgian NORM site)	
Indicators	Values
Root mean squared logarithmic error (RMSLE)	1.50×10^{-2}
BIAS	7.05×10^{-4}
Geometric mean bias (MG)	1.01
Cross-correlation function	Not correlated because the field situation is at near-equilibrium so not enough experimental data points to do the test.

Table 12.1-10: Qualitative performance assessment of ECOFOR applied to the Belgian NORM site

Qualitative performance table for ECOFOR (applied to the Belgian NORM site)	
Indicators	Properties
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> • 61 generic radionuclide-independent parameters covering the soil hydrology, vegetation, evapotranspiration, atmospheric and radionuclide-dependent information as well as general physical constants. • 113 site-specific data, including radionuclide-dependent (K_ds, plant selectivity coefficients, initial values) and radionuclide-independent for the 10 soil layers (diffusion rates, bulk and particle densities, fraction of sand and clay, hydraulic conductivities, field capacities and residual water content).
Number of processes considered separately	<ul style="list-style-type: none"> • Hydrological processes: water infiltration from rainfall, simplified representation of the balance between hygroscopic, capillary and gravitational water (Darcy flow, Lucas-Washburn capillary flow, tipping bucket switch-enabled model in adjacent soil layers). • Element transport processes in soil (linkage of elements to water transport by retardation, balance between exchangeable and unavailable soil fractions). • Uptake and transport processes for water and radionuclides: Root uptake, wilting point, anaerobiosis, water fluxes balanced against evapotranspiration, sap flow, interception, washout, absorption, translocation & leaching, litterfall, root decomposition and selectivity coefficients to link radionuclide transport to water fluxes in plants.
Application of the model to past set of data	<ul style="list-style-type: none"> • The model was previously applied to the Scots pine forest situated within the SCK-CEN domain, for which a monitoring station was purposely-build.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> • The previous model version for the Mol forest was not for radionuclides but for stable elements: Cl, Ca, K, Mg and Mn. The soil and element-dependent parameters were different, but the model had the same structure and processes.

12.1.2.2.2 Performance assessment analysis of TALAL model applied to Belgian NORM site

Table 12.1-11: Quantitative performance assessment of **TALAL** applied to the Belgian NORM site

Quantitative performance table for TALAL (applied to the Belgian NORM site)	
Indicators	Values⁹
Root mean squared logarithmic error (RMSLE)	<ul style="list-style-type: none"> • Uranium: 0.04 • Lead: 0.11 • Barium¹⁰: 0.38 all: 9.80E-02
BIAS	<ul style="list-style-type: none"> • Uranium: -0.002 • Lead: -0.009 • Barium: 0.18 all: 2.37E-03
Geometric mean bias (MG)	<ul style="list-style-type: none"> • Uranium: 1.11 • Lead: 1.28 • Barium: 0.42 all: 1.04
Cross-correlation function	Not correlated because the field situation is at near-equilibrium so not enough experimental data points to do the test.

⁹ The values are based on a single measurement-prediction comparison (i.e. one measurement compared to a single prediction) and using default parameter values (no calibration due to lack of data).

¹⁰ Barium is used as a stable surrogate for radium.

Table 12.1-12: Qualitative performance assessment of TALAL applied to the Belgian NORM site

Qualitative performance table for TALAL (applied to the Belgian NORM site)	
Indicators	Properties
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> • The number depends on what processes the user wants to simulate (e.g. soil moisture dynamics and water uptake by plants, solute dynamics and their uptake by plants, cycling of solutes in plant, feedback loops from plant to soil, etc.). • The model has maximum of 50 parameters.
Number of processes considered separately	<ul style="list-style-type: none"> • Water flow and uptake in soil-plant systems: <ul style="list-style-type: none"> ○ Rainfall interception by plant foliage and infiltration into soil, ○ Evaporation from soil surface, ○ Plant water uptake with adjustment depending on soil moisture (uptake is reduced when soil water available to plants decreases), ○ Percolation and deep drainage, ○ Redistribution of soil moisture between (unlimited) soil horizons, ○ Capillary rise from a shallow water table. • Solute transport and uptake in soil-plant systems: <ul style="list-style-type: none"> ○ Transport by advection, ○ Transport by dispersion (molecular diffusion and hydrodynamic dispersion), ○ Leaching into groundwater, ○ Uptake by plant roots according to linearised Michaelis-Menten kinetics, ○ Feedback loops between soil and plant compartments (disintegration of plant roots and litter layer). • Solute recycling in plants: <ul style="list-style-type: none"> ○ Exchange of solutes between three compartments: roots, trunk and foliage (number of compartments is adjustable)
Application of the model to past set of data	<ul style="list-style-type: none"> • The model is being validated using data from the experimental Scots pine forest near SCK-CEN.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> • The model is being validated using stable chlorine data from the experimental Scots pine forest near SCK-CEN.

12.1.3 FUKUSHIMA

12.1.3.1 Widely applied models

12.1.3.1.1 Performance assessment analysis of FORESTCROM model applied to FUKUSHIMA

Quantitative performance assessment: Makes use of available experimental data.

Qualitative performance assessment: Complements quantitative indicators; Provides indication about acceptability of a model in case data are not available to calculate any quantitative indicator.

Table 12.1-13: Quantitative performance assessment table for quantifying performance of FORESTCROM applied in the Fukushima (KU1-S)

Quantitative performance table for FORESTCROM (applied to the Fukushima site)	
Indicators	Values
Root mean squared logarithmic error (RMSLE)	Needles 0.22
	Wood 0.16
	Soil 0.67
RMSLE (aggregated)	0.72
BIAS	Needles -0.636 Bq/m ²
	Wood +42.8 Bq/kg
	Soil -6.4E03 Bq/kg
Geometric mean bias (MG)	Needles 0.79
	Wood 0.81
	Soil 4.39
MG aggregated	2.8
Cross-correlation function	There are no time series in the database

Table 12.1-14: Qualitative performance assessment performance table for quantifying performance of FORESTCROM applied in the Fukushima (KU1-S)

Qualitative performance table for FORESTCROM (applied to the Fukushima site)	
Indicators	Properties
Number of parameters involved and their time-independence	10
Number of processes considered separately	3
Application of the model to past set of data	The model was developed for Territories project as a demonstration
Application of the model to stable isotopes and chemical analogues	The model was developed for Territories project as a demonstration

12.1.3.1.2 Performance assessment analysis of TREE4-simple model applied to FUKUSHIMA

Table 12.1-15: Quantitative performance assessment of TREE4-simple applied to the Fukushima site

Quantitative performance table for TREE4-simple (applied to the Fukushima site)	
Indicators	Median values [5%, 95%]
Root mean squared error (RMSE) (n.d.)	<ul style="list-style-type: none"> Soil/tree partitioning: 0.055 [0.039, 0.077]
Root mean squared logarithmic error (RMSLE)	<ul style="list-style-type: none"> Concentrations in tree organs: 0.551 [0.441, 0.676] Concentrations in soil layers: 0.179 [0.144, 0.275] Depuration fluxes: 1.205 [0.771, 1.947]
BIAS (n.d.)	<ul style="list-style-type: none"> Soil/tree partitioning: -0.0013 [-0.0018, -0.0009]
Geometric mean bias (MG)	<ul style="list-style-type: none"> Concentrations in tree organs: 0.849 [0.574, 1.241] Concentrations in soil layers: 1.097 [0.781, 1.497] Depuration fluxes: 3.284 [1.729, 7.652]
Cross-correlation function	Cannot be calculated because too many gaps in the data time series.

Table 12.1-16: Qualitative performance assessment of TREE4-simple applied to the Fukushima site

Qualitative performance table for TREE4-simple (applied to the Fukushima site)	
Indicators	Properties
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> 22 “radiological” parameters involved in the calculation of radionuclide transfers in the forest stand, including 15 chemical-dependent parameters
Number of processes considered separately	<ul style="list-style-type: none"> 14 radionuclide transfer processes and 7 compartments. For further details, see the conceptual model of TREE4-simple and the associated description.
Application of the model to past set of data	<ul style="list-style-type: none"> The model was developed after Chernobyl accident and applied to European forests in the frame of EC and IAEA programs.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> The model is not recommended for stable analogues.

12.1.3.2 More advanced models

12.1.3.2.1 Performance assessment analysis of TREE4-advanced model applied to FUKUSHIMA

Table 12.1-17: Quantitative performance assessment of TREE4-advanced applied to the Fukushima site

Quantitative performance table for TREE4-advanced (applied to the Fukushima site)	
Indicators	Median values [5%, 95%]
Root mean squared error (RMSE) (n.d.)	<ul style="list-style-type: none"> Soil/tree partitioning 0.036 [0.030, 0.050]
Root mean squared logarithmic error (RMSLE)	<ul style="list-style-type: none"> Concentrations in tree organs: 0.281 [0.238, 0.347] Concentrations in soil layers: 0.189 [0.157, 0.345] Depuration fluxes: 0.575 [0.523, 0.640]
BIAS (n.d.)	<ul style="list-style-type: none"> Soil/tree partitioning: -0.008 [-0.11, -0.006]
Geometric mean bias (MG)	<ul style="list-style-type: none"> Concentrations in tree organs: 0.923 [0.673, 1.243] Concentrations in soil layers: 1.4 [0.83, 3.452] Depuration fluxes: 1.602 [1.253, 2.058]
Cross-correlation function	Cannot be calculated because too many gaps in the data time series.

Table 12.1-18: Qualitative performance assessment of TREE4-advanced applied to the Fukushima site

Qualitative performance table for TREE4-advanced (applied to the Fukushima site)	
Indicators	Properties
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> 34 “ecophysiological” parameters involved in the calculation of the forest stand growth, biomasses/area indexes of the tree organs and biomass fluxes. 6 “hydrological” parameters involved in the calculation of water fluxes in the forest stand. 36 “radiological” parameters involved in the calculation of radionuclide transfers in the forest stand, including 21 chemical-dependent parameters
Number of processes considered separately	<ul style="list-style-type: none"> 29 radionuclide transfer processes and 10 compartments. For further details, see the conceptual model of TREE4-advanced and the associated description.
Application of the model to past set of data	<ul style="list-style-type: none"> The model was developed after Fukushima accident and applied only to Japanese forests.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> The model can in principle be applied to stable isotopes or chemical analogs because its structure and the processes which are accounted for remain adequate, but the chemical-dependent parameters must be re-estimated.

12.1.4 SELLAFIELD

12.1.4.1 Widely applied models

12.1.4.1.1 Performance assessment analysis of the model ARCTIMAR applied to the Cumbrian Water region

Table 12.1-19: Quantitative performance assessment of ARCTIMAR applied to the Cumbrian Water region

Quantitative performance assessment of ARCTIMAR model applied to applied to the Cumbrian Water region including the intertidal beach region near the Sellafield nuclear facilities			
Medium	Values of indicators		
	Root mean squared logarithmic error (RMSLE)	BIAS	Geometric mean bias (MG)
filtered water			
Am-241	0.54	-2.63 Bq/m ³	3.25
Cs-137	0.58	5.30·10 ³ Bq/m ³	3.50
sediment			
Am-241	0.31	5.44·10 ² Bq/kg	0.51
Cs-137	0.23	1.01·10 ² Bq/kg	1.05
water and sediment			
Am-241	0.43		1.14
Cs-137	0.50		1.51
all data	0.44		1.35

Table 12.1-20: Qualitative performance assessment of ARCTIMAR applied to the Cumbrian Water region

Qualitative performance table for ARCTIMAR applied to the Cumbrian Water region	
Indicators	Values
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> The ARCTIMAR model was developed for radioecological assessment for the Norwegian Radiation Protection Authority (the NRPA box model) for the Arctic and North Seas. The model has various modifications. The model content about 2000 parameters. Although the model is dynamic, all the parameters are independent of the time.
Number of processes considered separately	<ul style="list-style-type: none"> Two main processes are considered in the model: (i) dispersion in the marine environment and (ii) bioaccumulation of radionuclides by biota and humans. The model includes advection of radioactivity between compartments, sedimentation, diffusion of radioactivity through pore water in sediments, particle mixing, remobilization, pore water mixing. a burial process of radioactivity in deep sediment layers and radioactive decay. Accumulation of contamination by biota is calculated from radionuclide concentrations in filtered seawater. Doses to humans are calculated on the basis from seafood consumptions. Dose rates to biota are developed on the basis of calculated radionuclide concentrations in marine organisms, water and sediment.
Application of the model to past set of data	<ul style="list-style-type: none"> The model has been used in many international projects. Simulations have been compared with different data sets and results from other model.
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> The model can be adjusted to stable elements.

12.1.4.2 More advanced models

12.1.4.2.1 Performance assessment analysis of AMIS model applied to applied to the Cumbrian Water region including the intertidal beach region near the Sellafield nuclear facilities

Table 12.1-21: Quantitative performance assessment of *AMIS* model applied to applied to the Cumbrian Water region including the intertidal beach region near the *Sellafield* nuclear facilities

Quantitative performance assessment of <i>AMIS</i> model applied to applied to the Cumbrian Water region including the intertidal beach region near the <i>Sellafield</i> nuclear facilities			
Medium	Values of indicators		
	Root mean squared logarithmic error (RMSLE)	BIAS	Geometric mean bias (MG)
filtered water			
Am-241	0.16	-0.46 Bq/m ³	0.92
Cs-137	0.27	5.35·10 ² Bq/m ³	1.94
sediment			
Am-241	0.15	1.35·10 ² Bq/kg	0.86
Cs-137	0.22	8.37·10 ¹ Bq/kg	1.06
water and sediment			
Am-241	0.15		0.88
Cs-137	0.26		1.26
total activity for the intertidal beach region			
Am-241	0.084	1.19·10 ⁵ Bq	0.95
Cs-137	0.30	-1.22·10 ⁵ Bq	0.86
all data	0.24		1.05

Table 12.1-22: Qualitative performance assessment of *AMIS* model applied to applied to the Cumbrian Water region including the intertidal beach region near the *Sellafield* nuclear facilities

Qualitative performance assessment of <i>AMIS</i> model applied to applied to the Cumbrian Water region including the intertidal beach region near the <i>Sellafield</i> nuclear facilities	
Indicators	Properties
Number of parameters involved and their time-independence	<ul style="list-style-type: none"> The model has about 200 parameters regarding the modelling processes and site-specific data
Number of processes considered separately	<ul style="list-style-type: none"> Parameters describing modelling processes presented in Table 6.4.1.2 are selected with regards to the site-specific data. Stokes law are used for preliminary estimation of the settlement process for the particles and potential release scenarios.
Application of the model to past set of data	<ul style="list-style-type: none"> The model was applied to the data-base, prepared in the course of the TERRITORIES project by Justin Smith <i>et al.</i> (2019) (Smith J. <i>et al.</i> D 9.59 - Radiological state database of sites. CONCERT- TERRITORIES Deliverable report, 2019)
Application of the model to stable isotopes and chemical analogues	<ul style="list-style-type: none"> the model is not intended for stable elements, but can be adjusted to stable elements.

12.2 Uncertainty assessment (Uncertainty matrix, Model Pedigree)

Uncertainty assessment comprises a range of methods ranging from identification of the various types of uncertainty to numerical quantification of their impacts on model outputs and sensitivity. This section provides a review of possible methodologies and potential applications within environmental modelling providing a supplement to the general guidance provided in the earlier deliverable D9.62 (Urso *et al.*, 2019) but specifically in relation to the model development theme elaborated in the current report.

12.2.1 The dimensions of uncertainty

In order to establish that a model is fit-for-purpose and the appropriate level of complexity has been selected, there may be a requirement that the uncertainties, which are inherent within the modelling process, are characterised and documented.

A useful way of structuring this has been presented by (Walker 2003) wherein 3 ‘dimensions’ of uncertainty have been defined as

- Location – the position within the assessment to which the uncertainty can be attributed. For example, the location could be placed at stages such as the problem formulation, the assessment context, the modelling inputs, the model structure and parameters, the model output interpretation and communication.
- Level – this encompasses the notion of a sliding scale or spectrum of uncertainty from complete determinism or mechanistic understanding (with the concomitant unattainable goal of being able to characterise in precise detail all germane aspects of a system) to complete ignorance extending even to the idea of unknown-unknowns (Kim, 2012).
- Nature – essentially split into ‘Knowledge-related’ (epistemic or epistemological) uncertainty defined as lack of scientific knowledge about specific factors, scenarios, parameters or models and ‘Variability’ (ontic/ontological or aleatory uncertainty) defined as the natural variability due to diversity or true heterogeneity in a data set or population.

Further dimensions have been suggested by (van der Sluijs 2006) on

- the qualification of knowledge base, which aims to map and characterise the underpinning elements thus enabling the identification of weak and strong parts in the assessment and
- value-ladenness of choices, which refers to the presence of values and biases in the various choices involved. This might, for example, pertain to choices concerning which:
 - scientific questions are posed,
 - data are selected, interpreted and rejected,
 - methodologies and models are devised and used,
 - explanations and conclusions are formulated etc.

Finally, the above categories may be extended to include different types of societal uncertainties, such as decisions about the acceptability of the risk, communication of the data, and variabilities in viewpoints between actors. This is particularly relevant for environmental risk, and debates on scientific uncertainty linked to climate change are a classic example (Strand and Oughton, 2009). These have obvious links to the value-ladenness category, but societal aspects can influence other dimensions. For example, linguistic uncertainty (uncertainty in language) might be included as a subcategory of uncertainty nature. This type of uncertainty arises because much of our natural language, including a great deal of our scientific vocabulary, is unspecific, ambiguous, vague, context dependent, or exhibits theoretical indeterminacies (Regan *et al.*, 2002).

This combined ‘typology’ and the subcategories of uncertainty that exist under each dimension were illustrated by (Zinger 2007) and are presented in Figure 12.2-1 below.

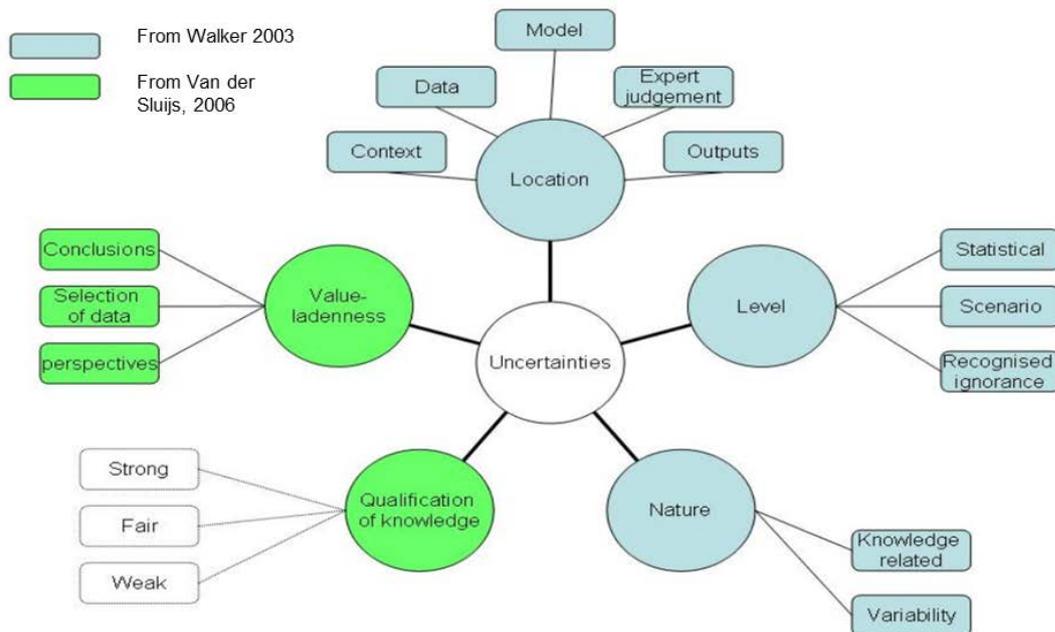


Figure 12.2-1: Illustration of the categorisation or ‘dimensions’ of uncertainty from (Zinger 2007) referring to the publications (Walker 2003) and (van der Sluijs 2006)

It should be acknowledged that not all uncertainties are quantifiable/numerical. There is often a tendency to focus on, for example, input and parameter uncertainties for which underlying statistical distributions are available, but these may not actually constitute the most important component of the underlying modelling uncertainty. Tools are available to document systematically these various types of uncertainty, e.g. NUSAP and uncertainty matrix, as discussed below.

Where uncertainties associated with any given model have been reported in published material, this has often been considered at the end of the procedure. However, a more efficacious approach may be to consider uncertainties at all stages of the assessment so that such an analysis becomes comprehensive and can be traced and audited throughout the modelling work.

12.2.2 Approaches available for uncertainty assessment

An overview of many of the approaches used in uncertainty assessment are covered by Urso *et al.* (2019) and the reader is referred to that report for an in-depth overview of the subject. The text below alludes to some of this work with the addition of supplementary material where this was deemed useful.

Baustert *et al.* (2018) provide an overview of those frameworks/guidance efforts that the authors deemed valuable to advance the practice of Uncertainty Assessment (UA), seen from the perspective of integrated environmental models for ecosystem service assessments. By way of example, one of the frameworks/guidance efforts has been selected for more detailed consideration, that of Refsgaard *et al.* (2007), to illustrate some of the typical elements that may be included in UA. Refsgaard *et al.* (2007) provide 14 methods that may be applied and these are listed as:

- Uncertainty matrix [considered by Urso *et al.* (2019) but additional consideration provided here]
- NUSAP [briefly mentioned by Urso *et al.* (2019) but additional consideration provided here]
- Data uncertainty engine
- Error propagation equations [considered by Urso *et al.* (2019) but additional consideration provided here]
- Extended peer review (review by stakeholders) (cf. TERRITORIES WP3, as introduced by Guillevic *et al.* (2018))
- Inverse modelling (parameter estimation)
- Inverse modelling (predictive uncertainty)
- Monte Carlo analysis (Urso *et al.*, 2019)
- Multiple model simulation
- Quality assurance [considered above]
- Scenario analysis [Scenario uncertainty covered by Urso *et al.* (2019)]
- Sensitivity analysis [covered by Urso *et al.* (2019)]
- Stakeholder involvement [cf. TERRITORIES WP3, as introduced by Guillevic *et al.* (2018)]
- Expert elicitation [considered by Urso *et al.* (2019) but additional consideration provided here]

Some of these methods are described elsewhere in the TERRITORIES project as noted in the above list and further elaboration is considered unnecessary here. So, for example Monte Carlo analysis, often forming a component of uncertainty analysis, has been described in great detail, as has sensitivity analysis, in TERRITORIES Deliverable Report D.9.62 (Urso *et al.*, 2019) and further, repetitive analysis here was considered unnecessary, Stakeholder involvement has formed the central theme for TERRITORIES WP3 (as introduced by Guillevic *et al.*, 2018), Quality assurance is covered in Section 9.1.3. This then leaves several other methods that are elaborated below. The overview starts with methods of categorising uncertainties, then progresses to more qualitative and quantitative methods that consider the quality and robustness of the data as well as numerical data handling before ending with expert elicitation.

Uncertainty matrix

As described by Walker *et al.* (2003), the purpose of an uncertainty matrix (Table 12.2-1) is to provide a tool by which to attain a systematic and graphical overview of the essential features of uncertainty in relation to the use of models in decision support activities. The approach is essentially a means of mapping and documenting the dimensions of uncertainty described in Section 12.2.1.

Table 12.2-1: The uncertainty matrix – reproduced from Walker *et al.* (2003).

Location		Level			Nature	
		<i>Statistical uncertainty</i>	<i>Scenario uncertainty</i>	<i>Recognised ignorance</i>	<i>Epistemic uncertainty</i>	<i>Variability uncertainty</i>
Context	Natural, technological, economic, social and political representation					
Model	Model structure					
	Technical model					
Inputs	Driving forces					
	System data					
Parameters						
Model Outcomes						

The uncertainty matrix approach has not been applied formally in WP1 models but as was the case for scenario analysis, the authors of this report are familiar with the approach having been involved in its application in other circumstances. For example, the uncertainty matrix was applied in the development of the ERICA Tool as described by Zinger (2007) and Oughton *et al.* (2008) as illustrated in Table 12.2-2.

Table 12.2-2: Example of application of the uncertainty matrix to a component (Concentration ratios: CRs in the Assessment Tool) of the ERICA integrated approach; reproduced from Oughton *et al.* (2008).

Location		Type of uncertainty			Nature of uncertainty		
		Statistical	Scenario - range	Ignorance	Knowledge-related	Inherent variability	Quality of knowledge base
Assessment Tool							
Model Parameters	CRs	A : Site specific CRs (e.g. in Tier 3)			Conceptual and model uncertainties related to the use of simple equilibrium factors to model complex dynamic processes – applies to any use of CRs	Dependent on the amount of data available. Variability may be underestimated if sampling is not spatially or temporarily representative	Good – specific to situation being considered (although question whether site specific data are always best)
		B : Generic data for Cs-137 and Sr-90 distribution data and statistics available	Applicability of generic data to the case and scenario under assessment	May be gaps for protected or rare species	As above	Large amount of data and appropriate sampling and analysis. Site-specific applicability may be unknown.	Good for many organisms. However, there will be gaps for some organisms and life history stages as some of the CR database is related to human food-chain modelling
			C : Choice of CRs based on expert judgement and extrapolation methods, e.g. on trace or similar elements		As above	Significant – related to site-specific variation and variations in radionuclide/organism characteristics	Depends on radionuclide and organism involved – may vary between moderate and poor.
				D : No data available for radionuclides or organisms or life history stages – apply maximizing assumptions	Applicability of the level of conservatism assumed	No information. Use of assumed probability distributions may be required	Poor knowledge base

The use of uncertainty matrix has been established as being a useful tool in the mapping of uncertainties associated with (the characterisation and modelling of) complex socio-ecological systems as evidenced by recent applications with environmental sciences, such as the environmental management of the Baltic Sea (Udovyk and Gilek, 2013). Nonetheless, it is sobering to note that these authors, when considering the broader availability of theoretical approaches (including but not limited to uncertainty matrix) for coping with uncertainty in complex socio-ecological systems, concluded that although these were ample, a challenge remained in implementing these approaches more effectively in assessment and management frameworks.

NUSAP

NUSAP was briefly mentioned by Urso *et al.* (2019) and is further developed here. It builds on work by Walker and Van der Sluis, and extends the mapping in uncertainty matrixes to a more systematic assessment of the quality of data. The NUSAP acronym stands for numeral, unit, spread, assessment,

and pedigree. ‘Numeral’ is expressed as a measure of a quantity, such as an arithmetic mean of a given parameter/endpoint, and expressed with a physical ‘unit’ such as kilogram (kg). This category also allows for the inclusion of additional information such as location where or time when the quantity was assessed/evaluated. The ‘Spread’ category provides a quantitative representation of uncertainty using, for example, the range in possible values (min, max) or statistical methods to create intervals to characterise the variability/inexactness of the quantity/information in the Numeral qualifier. The last 2 qualifiers in the NUSAP system, cover other aspects of model uncertainty with regards to ‘assessment’, which incorporates the concepts of expert judgement of (model) reliability and ‘pedigree’, which is used to determine the robustness of different phases of the knowledge base production. In this regard, Pedigree criteria can include: empirical basis, methodological rigor, proxy representation, theoretical understanding, and degree of validation. Pedigree assessment can be further extended to also address societal dimensions of uncertainty, using criteria that address different types of value ladenness, quality of problem frames (Refsgaard *et al.*, 2007). An example is given in Table 12.2-3.

Table 12.2-3: Pedigree matrix for reviewing the knowledge base of assumptions (from Van der Sluijs *et al.*, 2005)

Criterion	Score		
	2	1	0
Plausibility	Plausible	Acceptable	Fictive or speculative
Inter-subjectivity peers	Many would make same assumption	Several would make same assumption	Few would make same assumption
Inter-subjectivity stakeholders	Many would make same assumption	Several would make same assumption	Few would make same assumption
Choice space	Hardly any alternative assumptions available	Limited choice from alternative assumptions	Ample choice from alternative assumptions
Influence situational limitations (time, money, etc.)	Choice assumption hardly influenced	Choice assumption moderately influenced	Totally different assumption when no limitations
Sensitivity to view and interests of the analyst	Choice assumption hardly sensitive	Choice assumption moderately sensitive	Choice assumption sensitive
Influence on results	Only local influence	Greatly determines the results of link in chain	Greatly determines the results of the indicator

NUSAP was designed to provide an analysis and diagnosis of uncertainty in science for policy. It has been widely applied, recent examples being those involving modelling for policy support in the UK (Pye *et al.*, 2018) and beach quality indexes in coastal management (Bombana & Ariza, 2018). Moreover, there has been recognition recently in other scientific sectors, such as those associated with the oil and gas sector, that efforts should be made to characterise uncertainties that are not amendable to quantification, such as those related to the reliability and robustness of underpinning scientific knowledge. Berner & Flage (2016) referred to this as ‘Extended quantitative risk assessments’ and concluded that there were strong parallels with the NUSAP approach and that the use of visualisation tools in NUSAP were of particular utility. It should be noted, therefore, that NUSAP is one example of many ‘Uncertainty management Techniques’, an insightful overview being provided by Skinner *et al.* (2017) in relation to environmental risk assessments generally.

Data uncertainty engine (DUE)

Brown *et al.* (2005) provide a framework for recording uncertainties about environmental data that is both broadly applicable to different types of data and adaptable for different complexities of problem. The conceptual model for recording uncertainties about environmental data is based on a fourfold distinction between: 1) the empirical quality of data; 2) the sources of uncertainty in data; 3) estimates of ‘fitness for use’, where information about specific problems that have been, or might be, encountered in applying these data are recorded; and 4) the ‘goodness’ of an uncertainty model.

Of particular note here is the category ‘fitness-for-use’ as used by Brown *et al.* (2005) which is clearly closely related to the subject of this report on the ‘fit-for-purpose level of complexity’. This comprises a description of the application(s) for which the data are intended (or have been used) and of the problems associated with those applications.

Error propagation equations

The analytical approach to propagate uncertainty (errors) was introduced in section 5.2 of CONCERT-TERRITORIES deliverable report D9.62 (Urso *et al.*, 2019). It is further developed here.

Errors can be propagated through a simple system, for example those involving basic arithmetic operations of independent variables (x and y below), to yield an overall Standard Error, SE, via simple equations of the form:

$$SE(x + y) = \sqrt{SE(x)^2 + SE(y)^2} \quad (12.1)$$

$$SE(x - y) = \sqrt{SE(x)^2 + SE(y)^2} \quad (12.2)$$

When adding and subtracting variables

$$SE(x \cdot y) = x \cdot y \sqrt{\left[\frac{SE(x)}{x} \right]^2 + \left[\frac{SE(y)}{y} \right]^2} \quad (12.3)$$

$$SE(x/y) = (x/y) \sqrt{\left[\frac{SE(x)}{x} \right]^2 + \left[\frac{SE(y)}{y} \right]^2} \quad (12.4)$$

When multiplying and dividing variables.

However, there are numerous assumptions made including those associated with the normality of underlying distributions associated with each variable, that they have been sampled independently and that there is no covariance between the variables. These assumptions rarely hold and the application of this method is, in practice, restricted to simple systems preliminary screening analysis.

Inverse modelling (parameter estimation & predictive uncertainty)

As noted by Refsgaard *et al.* (2007), parameter values can be estimated through inverse modelling. For simple systems, this can be achieved by minimising an objective function, i.e. minimise the difference between the summed squared deviation between the calibration endpoints (e.g. field data) and their modelled counterparts.

Within the scientific discipline involving the study of environmental radioactivity a commonly encountered challenge relates to the characterisation of parameters associated with a given source term. Examples include the work of Pudykiewicz (1998) who used an adjoint tracer transport equation to evaluate the emission field of atmospheric contaminants. This enabled important parameters (notably the source strength and the location) to be established. (Penenko *et al.*, 2002) used inverse (radionuclide atmospheric dispersion) modelling techniques, for source parameters estimations.

Optimisation routines often exist in modelling platforms, once more providing ECOLEGO (Avila *et al.*, 2005) as an example, and within bespoke models, *per se*, such as the segmented Gaussian plume model (SGPM) where a optimisation approach for recursive parameter estimation was developed by (Pecha & Smídl, 2016). Refsgaard *et al.* (2007) emphasise that some of the inverse optimisation routines include the ability to estimate predictive uncertainties. Although the method provides an objective estimate of the predictive uncertainty, given the applied model structure, the approach is limited by assumptions concerning linearity and normally distributed residuals and by the fact that uncertainty can only be predicted for data types for which observations exist.

Multiple model simulation

As the name suggest, the multiple model simulation approach involves the application of many different models to the same problem formulation and environmental conditions. In this way insights are gained into model structural uncertainty. As considered by Refsgaard *et al.* (2007), an important limitation pertains to whether the relevant space of plausible models has been adequately sampled and whether important plausible model structures have been overlooked. Structural model uncertainty can be greater than one might expect. For example, Engeland *et al.* (2005) showed that the effect of the model structural uncertainty on the total simulation uncertainty for a conceptual water balance model was greater than parameter uncertainty. This approach was generally not applied in TERRITORIES although multiple models were applied at the Fen Site (Section 5) for the sake of comparing the efficacy of model prediction for dose rates in air. No efforts were made, however, to quantify model structural uncertainty although it is recognised that such analysis may have provided useful insight.

Scenario analysis

Scenario uncertainty has been introduced by Urso *et al.* (2019) for the radioecological scenario, and by Jones *et al.* (2019) for the exposure scenario.

Starting from the premise that if the future is uncertain Van Der Heijden (2000) contended that there are, in fact, multiple equally plausible futures, and that these can be referred to as scenarios. From this perspective, a total set of scenarios about the future reflects our understanding of what in the system is predetermined and predictable, and what we believe to be fundamentally indeterminate (Van Der Heijden, 2000). An advantage in the application of Scenarios is that the approach can ensure that

assumptions about future developments are made transparent and documented and are often the only way to deal with an unknown future.

According to Alcamo (2001), Scenarios can be assigned to 3 major classes consisting of opposing types these being:

- deductive vs. inductive. Deductive scenarios are derived from a framework which organises the big questions or uncertainties about the future into a logical form. Initially the framework is established and thereafter scenarios are deduced from the framework. In contrast, inductive scenarios are derived from considering all data and ideas about the future. Scenarios are built step-wise and from the bottom-up.
- exploratory vs. anticipatory. Exploratory/descriptive scenarios start from the current situation and then describe the steps that lead to a future situation. On the other hand, anticipatory scenarios begin with a prescribed vision of the future (optimistic, pessimistic, or neutral) and then work backwards in time to visualise how this future could emerge.
- qualitative vs. quantitative. Qualitative Scenarios are either in the form of visual aids: e.g. video clips, or in the form of words: e.g. written phrases, outlines or storylines. The most common form is a storyline: i.e. narrative description of scenario, highlighting, for example, the main features of a case. Quantitative scenarios deal with numerical information and are commonly computed with models. Quantitative scenarios provide numerical information with identifiable underlying assumptions, but the models used have limited view of the world and are often not transparent. The exactness of model output may give a misleading illusion of certainty. A limitation for qualitative scenarios is that it is difficult to test the underlying assumptions. For quantitative scenarios, the analysis is limited to those aspects of reality that can be quantified.

Scenario analysis was not applied as an approach explicitly within the TERRITORIES modelling applications under WP1. This primarily reflects the fact that the methodology, self-evidently, lends itself most appropriately to predictive modelling, where sensible statements need to be made about future developments and conditions (especially in the longer term), whereas this was not the main focus of radioecological modelling. The authors of this Deliverable Report, nonetheless have broad experience with the application of Scenario analysis. In the work of Brown *et al.*, (2016) and Hosseini *et al.* (2017), for example, a scenario based approach was used to consider the potential future impact of releases from radioactive sources in the Arctic seas. The approach was, using the categorisations given above, essentially exploratory and quantitative in nature and involved, inter alia, sophisticated computer codes to consider the steps leading to and the plausible releases from criticalities occurring in sunken nuclear reactors.

Expert elicitation

Expert elicitation was briefly introduced in section 7 of CONCERT-TERRITORIES deliverable report D9.62 (Urso *et al.*, 2019). It is further developed here.

In some cases, the coverage of information pertaining to a given assessment (e.g. modelling inputs and/or parameters) is limited and some additional means of attaining data is required. This is arguably closely linked to stakeholder involvement but might be considered to be one component of that more general theme.

Taylor (1993) provides a useful overview of methods commonly used in exposure assessments to quantify incomplete knowledge, including informal and formal expert elicitation. Commonly asked questions might at the simplest level relate to the range of the data establishing what are the lowest

and highest possible values. More complex questions may involve establishing whether the distribution ought to be truncated and to the form of the distribution - is it skewed, flat or multi-modal? The author also draws attention to some of the cognitive biases (see for example Kahneman & Tversky, 1996) that may be involved in expert elicitation. Refsgaard et al. (2007) also note that elicitation protocols have been developed and the reader is referred to this publication should further information be required.

In some cases, empirical or mechanistically-underpinned datasets will be available and there may be some advantages in combing this information with that from expert elicitation. In this circumstance, suitable methods such as Bayesian approaches may be adopted as considered elsewhere (Hosseini et al., 2013).

12.2.3 Selection of appropriate methodology

The reader is referred to Urso *et al.* (2019) for general Guidance on the application of uncertainty analysis for radioecological models. Some further points are noted below.

As can be seen from the above overview, the number of available approaches for characterising uncertainty are not restrictive and it is arguably more demanding to select the correct method(s) for any given model development and application. The guidance would not be to apply every one of the above-mentioned methods indiscriminately but rather to tailor the selection so that it may be used more efficaciously to establishing whether the model has an appropriate fit-for-purpose level of complexity. Fortunately, Refsgaard *et al.*, (2007) also provide a very useful guidance in this regard by considering which methodology should be selected for different purposes and in different situations from three different perspectives. The deliberations in this paper are not repeated exhaustively here, the reader is referred to Refsgaard *et al.*, (2007) for a more detailed explanation, but some of the more pertinent points have been extracted for consideration below.

Refsgaard *et al.*, (2007) note that the stages within the modelling process and level of ambition may be an important perspective for enlightening the selection procedure. Initial stages, during model planning, should at a minimum require sources of uncertainty to be identified and this could be achieved using an uncertainty matrix at a basic level but may require additional methods, such as stakeholder involvement should more comprehensive assessment be required. For the main stages of the modelling assessment (as considered in Section 9.1.2, Steps 2-5) when review dialogue and decisions are required, quality assurance may be a key tool for a basic level of ambition but may be augmented using extended peer review and an updated uncertainty matrix for more comprehensive assessments. Finally, at the latter stages of model development (as considered in Section 9.1.2, Steps 4-5) when uncertainty assessment and propagation is needed, many of the numerical methods listed above, for example, Monte Carlo simulation, inverse modelling and sensitivity analysis, may be apposite supported using uncertainty assessment/mapping procedures such as NUSAP and Expert Elicitation. If the level of modelling ambition is lower, the list of applicable methods may be truncated to include methods such as data uncertainty engine, error propagation equations and sensitivity analysis. Refsgaard *et al.*, (2007) also structure the selection procedure from the perspective of purposes of use of uncertainty mapping methods. The most appropriate suite of methodologies differ depending on whether the objective is to identify sources of uncertainty, assess the levels of uncertainty for the various sources of uncertainty, propagate uncertainty through the model, trace and rank sources of uncertainty or to reduce uncertainty. The reader is referred to Refsgaard *et al.*, (2007) for further details.

Concluding remarks:

It is not possible to be prescriptive, in terms of which approaches should be applied in addressing uncertainties associated with model development for (radiological) risk assessment, thereby informing the fit-for-purposiveness level of complexity in model selection, because the spectrum of plausible requirements and applications is open-ended. The experience gained directly from the TERRITORIES model development and application exercise has been limited with only some examples, such as Monte Carlo analysis, Multi-model simulation and sensitivity analyses being provided on an *ad hoc* basis. Nonetheless, a general recommendation is that efforts should be made to map and characterise uncertainties at all stages of the model development where practicable, possibly using some of, but not restricting oneself to, the methods outlined above and those in CONCERT-TERRITORIES deliverable report D9.62 (Urso *et al.*, 2019). It is undoubtedly of great utility to define clearly, at the problem formulation stage, the criteria against which the efficacy of the model will be judged and that this list might include information defining an acceptable level of uncertainty in model outputs. Uncertainty assessment will ideally constitute an important intrinsic component in establishing fit-for-purposiveness and the appropriate level of complexity requirements for a given model.

12.2.4 References for Appendix 12.2

- Alcamo, J. (2001). Scenarios as Tools for International Environmental Assessments. European Environment Agency, Copenhagen. Environmental Issues Report. Experts Corner Report. Prospects and Scenarios No. 5.
- Avila R, Broed R, Pereira A (2005) ECOLEGO - A toolbox for radioecological risk assessment. In: Proceedings of the International Conference on the Protection from the Effects of Ionizing Radiation, Stockholm (Sweden) 6.-10. October 2003. IAEA-CN-109/80. Vienna: International Atomic Energy Agency, 2003: 229-232.
- Baustert, P., Othoniel, Benoit, Rugani, Benedetto, Leopold, Ulrich. (2018). Uncertainty analysis in integrated environmental models for ecosystem service assessments: Frameworks, challenges and gaps. *Ecosystem Services*, 33, 110-123. doi: 10.1016/j.ecoser.2018.08.007
- Berner, C. L., & Flage, R. (2016). Comparing and integrating the NUSAP notational scheme with an uncertainty based risk perspective. *Reliability Engineering & System Safety*, 156, 185-194. doi: 10.1016/j.ress.2016.08.001
- Bombana, B., & Ariza, E. (2018). Clarifying some assumptions of coastal management: Analysis of values and uncertainties embedded in beach quality indexes. *Ecological Indicators*, 91, 376-385. doi: 10.1016/j.ecolind.2018.03.066
- Brown, J. D., Heuvelink, G. B., & Refsgaard, J. C. (2005). An integrated methodology for recording uncertainties about environmental data. *Water Science and Technology*, 52(6), 153-160. doi: 10.2166/wst.2005.0163
- Brown, J. E., Amundsen, I., Bartnicki, J., Dowdall, M., Dyve, J. E., Hosseini, A., . . . Standring, W. (2016). Impacts on the terrestrial environment in case of a hypothetical accident involving the recovery of the dumped Russian submarine K-27. *J Environ Radioact*, 165, 1-12. doi: 10.1016/j.jenvrad.2016.08.015
- 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). doi: 10.1623/hysj.50.1.45.56334
- Hosseini, A., Stenberg, K., Avila, R., Beresford, N. A., & Brown, J. E. (2013). Application of the Bayesian approach for derivation of PDFs for concentration ratio values. *J Environ Radioact*, 126, 376-387. doi: 10.1016/j.jenvrad.2013.04.007

- Hosseini, A., Amundsen, I., Brown, J., Dowdall, M., Karcher, M., Kauker, F., & Schnur, R. (2017). Impacts on the marine environment in the case of a hypothetical accident involving the recovery of the dumped Russian submarine K-27, based on dispersion of (137)Cs. *J Environ Radioact*, 167, 170-179. doi: 10.1016/j.jenvrad.2016.11.032
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103 (No.3), 582-591.
- Kim, S. D. (2012). Characterizing unknown unknowns. Paper presented at the Paper presented at PMI® Global Congress 2012, North America, Vancouver, British Columbia, Canada.
- Oughton, D. H., Agüero, A., Avila, R., Brown, J. E., Copplestone, D., & Gilek, M. (2008). Addressing uncertainties in the ERICA Integrated Approach. *J Environ Radioact*, 99(9), 1384-1392. doi: 10.1016/j.jenvrad.2008.03.005
- Pecha, P., & Smidl, V. (2016). Inverse modelling for real-time estimation of radiological consequences in the early stage of an accidental radioactivity release. *J Environ Radioact*, 164, 377-394. doi: 10.1016/j.jenvrad.2016.06.016
- Penenko, V., Balanov, A., & Tsvetova, E. (2002). Methods of sensitivity theory and inverse modeling for estimation of source parameters. *Future Generation Computer Systems*, 18, 661-671.
- Pudykiewicz, J. A. (1998). Application of adjoint tracer transport equations for evaluating source parameters. *Atmospheric Environment*, 32 (No. 17), 3039-3050.
- Pye, S., Li, F. G. N., Petersen, A., Broad, O., McDowall, W., Price, J., & Usher, W. (2018). Assessing qualitative and quantitative dimensions of uncertainty in energy modelling for policy support in the United Kingdom. *Energy Research & Social Science*, 46, 332-344. doi: 10.1016/j.erss.2018.07.028
- Refsgaard, J. C., van der Sluijs, Jeroen P., Højberg, Anker Lajer, Vanrolleghem, Peter A. (2007). Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software*, 22(11), 1543-1556. doi: 10.1016/j.envsoft.2007.02.004
- Regan, H. M., COLYVAN M., & BURGMAN, M. A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*, 12(2) 618–628.
- Skinner, D. J. C., Rocks, S. A., & Pollard, S. J. T. (2017). Where do uncertainties reside within environmental risk assessments? Testing UnISERA, a guide for uncertainty assessment. *Environ Pollut*, 225, 390-402. doi: 10.1016/j.envpol.2017.02.065
- Strand, R, Oughton, D.H 2009 Risk and Uncertainty - as a Research Ethics Challenge. NENT, Oslo www.etikkom.no/Aktuelt/publikasjoner/Risk-and-Uncertainty---as-a-Research-Ethics-Challenge/
- Taylor, A. (1993). Using objective and subjective information to develop distributions for probabilistic exposure assessment. *Journal of Exposure Analysis and Environmental Epidemiology*, 3 (No.3), 285-298.
- Udovik, O., & Gilek, M. (2013). Coping with uncertainties in science-based advice informing environmental management of the Baltic Sea. *Environmental Science & Policy*, 29, 12-23. doi: 10.1016/j.envsci.2013.01.015.
- Urso, L. *et al.* (2019). D 9.62 – Guidance on uncertainty analysis for radioecological models. CONCERT-TERRITORIES Deliverable report.
- Van der Heijden, K. (2000). Scenarios and forecasting: Two perspectives. *Technological Forecasting and Social Change*, 65, 31–36.
- van der Sluijs, J. P. (2005). Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: The NUSAP System. *Risk Analysis*, 25 (No. 2), 481-492.
- van der Sluijs, J. P. (2006). Uncertainty, assumptions, and value commitments in the knowledge-base of complex environmental problems In S. G. V. a. S. Ângela Guimarães Pereira (Ed.), *Interfaces between Science and Society* (pp. 67-84.): Green Leaf Publishing.
- Walker, W. E., Haremoës P., Rotman, J., Van der Sluijs, J.P., Van asselt, M.B.A., Janssen P., and Kraymer Von Krauss, M.P. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment Vol. 4, No. 1*, 5–17.

Zinger, I., Copplestone, D., Brown, J., Sjöblom, K.L., Jones, S., Proehl, G., Oughton, D., Garnier-Laplace, J. (2007). Considerations for applying the ERICA Integrated Approach. Deliverable D8. EC project contract no. FI6R-CT-2004-508847.