Probabilistic and full-chain risk assessment of the chemical accumulation on human body using an integrated modelling tool

By

Taku Tanaka

UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI AGRARIA

DIPARTIMENTO DI SCIENZE E TECNOLOGIE AGROALIMENTARI E MICROBIOLOGICHE

CORSO DI DOTTORATO DI RICERCA IN
CHIMICA, BIOCHIMICA ED ECOLOGIA DEGLI ANTIPARASSITARI. CICLO XXIII

Matriculation number: R07760

Tutor: Prof. Ettore Capri

Coordinator: Prof. Claudia Sorlini

ANNO ACCADEMICO 2009/2010

Abstract

Mathematical models have been developed to address diverse issues. Especially environmental multimedia models are now well recognized as useful tools for environmental and health risk assessment and management. This dissertation consists of three main pillars: (i) the introduction of the newly developed integrated modelling tool (2-FUN tool) that deals with environmental and health risk assessment, (ii) the investigation of statistical approaches to derive density functions (PDFs) of input parameters of interest, and (iii) the application of the 2-FUN tool for a designed case study.

Chapter 1 presents general roles of mathematical models for health and environmental risk assessment of chemicals, the importance to consider parametric uncertainty, and the extensive review on existing modelling methodologies. At the end of the model review, main features of the 2-FUN tool are pointed out as follows: (i) Its capability to conduct a full-chain risk assessment on a common system, which allows linking the simulation of chemical fate in the environmental media, multiple pathways of exposure, and the detailed analysis for multiple effects in different target tissues in human body, (ii) it contains a wide range of methods for sensitivity and uncertainty analyses, and (iii) it can be user-friendly because of its effective graphical simulation interface and its flexibility, which facilitates users to design scenarios for target regions and arrange the tool on their own ways.

The conceptual and theoretical aspects of the 2-FUN tool are summarized in Chapter 2. The focus of Chapter 3 is set at the detailed presentation of two advanced statistical approaches to derive probabilistic density function (PDF) for the two parameters used in the freshwater compartment of the 2-FUN tool: the settling velocity of particles that is a driving factor influencing the transfer of particles at the water-sediment interface in fresh-water system, and the fish bioconcentration factor (BCF) for metal that represents the accumulation of a given chemical in organisms arising by water uptake.

Finally, Chapter 4 depicts how the 2-FUN tool can be applied based on a designed case study and shows the feasibility of the integrated modelling approach to couple an environmental multimedia and a PBPK models, considering multi-exposure pathways, and thus the potential applicability of the 2-FUN tool for health risk assessment.

Acknowledgement

TABLE OF CONTENTS

1 GENERAL INTRODUCTION

1.1 *Overview of environmental and health risk assessment*

Many of environmental problems caused by anthropogenic factors, i.e., increases in human population, energy consumption, agricultural land, and deforestation have been certainly increasing. According to the belief that our health and ecosystems are greatly influenced by such environmental problems, the efforts to control and regulate the environment become primary standards of public health policy throughout the world. The relationship between environmental problems and human health (and also ecological activities) has become one of major concerns of international, regional, and domestic organizations such as the World Health Organization (WHO), the European Environment Agency (EEA), the US Environmental Protection Agency (USEPA), and the Ministry Of Environment in Japan (MOE).

WHO issued the report describing the progress made by the WHO European Member States in improving their health and the environment situation over the last 20 years (WHO, 2010). Based on the Children Environment and Health Action Plan for Europe (CEHAPE) agreed in 2004, the assessment focuses on the aspects of health related to clean water and air, to environments supporting safe mobility and physical activity, chemical safety, noise and safety at work. Main conclusions drawn from the comprehensive assessment are summarized as follows;

- The improvement in the European health situation over the last two decades can be seen in many relevant issues, including better accessibility to improved water, reduced incidence of injuries, improved air quality and reduced exposure to lead and persistent organic pollutants.

- Despite the overall progress, significant disparities in health risks remain in the European region in relation to all priority issues listed in CEHAPE. These disparities are seen both between and within the countries. Several old issues remain unresolved, posing a significant public health problem in parts of the region, i.e., (i) poor access to improved water for the

rural population, (ii) population exposure to inhalable Particulate Matter (PM) remains stable after a substantial fall in the 1990s, (iii) 80 % of children are regularly exposed to Second-Hand tobacco Smoke (SHS) at home, and more than 20% of households live in houses subject to damp and mould, and (iv) many countries in Europe continue to use leaded petrol.

The key messages concerning agricultural ecosystems were posed by EEA (EEA 2010) as follows;

- Within the framework of the Common Agricultural Policy (CAP), the last 50 years have seen increasing attention to biodiversity, but without clear benefits up to present.
- With agriculture covering about half of EU land area, Europe's biodiversity is deeply linked to agricultural practices, creating valuable agro-ecosystems across the whole Europe.
- Biodiversity in agro-ecosystems is under considerable pressure as a result of intensified farming and land abandonment.

The management of the environment and any associated human health impacts using such science-based tools as toxicological evaluation, risk assessment, and economic evaluation has become widely recognized in professional circles. Environmental risk assessment (ERA) has been developing as a technique and a profession since the 1970s (Ball 2002). According to IPCS and OECD (2003), "risk assessment" is a process to calculate or estimate the risk to a target organism, system or population following exposure to an agent. It includes hazard identification, hazard characterization, exposure assessment, and risk characterization. The presence of variety of the concepts and definitions on "risk assessment" implies that risk assessments are given specific meanings in concrete uses.

"Integrated risk assessment" has been defined as "a process that combines risks from multiple sources, stressors, and routes of exposure for humans, biota, and ecological resources in one assessment" (USEPA 2002). Related concepts include cumulative, holistic and comparative risk assessment (Power and McCarty 1998). The cumulative risk assessment of USEPA (1997 and 2003) addresses mixtures of chemicals, long-term risks, and various endpoints. IPCS (2001) focused on integrating risks to human and non-human receptors. In the EU environment and health strategy, the concept is treated broadly, seeking to integrate information; research; environmental, health and other policies; cycles of pollutants; interventions; and stakeholders (CEC 2003). Table 1.1 presents key types of integration in risk assessment identified in IPCS (2001).

"Risk management" includes technical and non-technical aspects (IPCS and OECD 2003) in various stages from risk prevention, avoidance and reduction to compensation. The borderline between risk assessment and management is not clear (Assmuth and Hilden 2008). They are interacting and partly overlapping parts of a process. Management includes interventions to alter the risks and involves greater intentionality and associated value judgments than "pure assessments", for these interventions to take place. This means that the management frameworks, especially on a policy level, can be more loosely structured than the assessment frameworks and more open to a diverse set of information.

"Risk" has been defined as "the probability of an adverse effect in an organism, system or population caused under specified circumstances by exposure to an agent" (IPGS and OECD 2003). In quantitative analyses risk is defined as a function (usually the product) of the probability and consequences of an adverse event or process. A chain of events or stages of risk formation can be separated from exposures to effects. Risk concepts are used in a wide range of meanings, and different uses and users emphasize different aspects of risks (Assmuth and Hilden 2008).

Table 1.1: Key types of integration in risk assessment identified in IPCS (2001)

Integration dimension	Description and recommendation by IPCS (2001)
Agents	Integrate risks from all relevant agents, e.g., risks to aquatic life from pesticides are routinely considered but for the restoration in agricultural areas, silting, fertilization, and channel destruction should also be considered
Routes	Integrate risks from all relevant routes of exposure, e.g., assessments of pesticides may need to consider routes of exposure other than diet
Endpoints	Consider all relevant potentially significant endpoints; mechanistic understanding allow multiple endpoints to be assessed in a common and consistent manner; multiple endpoints may also be integrated into common units such as quality adjusted life years
Receptors	Consider all relevant classes of receptors, e.g., the entire exposed population, including all age classes, not just maximally exposed individuals, and the distribution of risks across the exposed biotic community
Scales	Extrapolations in risk assessment can occur in various dimensions including time, place, space, biological scale, or mechanisms; integration of various scales and dimensions may be necessary
Life cycle stages	May need to integrate the risks from the entire life cycle of a chemical or product, including production of raw materials, manufacturing, use, and

	disposal of both the product and associated byproducts
Management options	When decisions are based on comparison of alternatives, assessments should consider the risks from relevant alternatives in an integrated manner; e.g., an assessment of risks from waste water should consider both the risks from the untreated effluent and the risks from alternative treatment technologies including disposal of sludge
Socioeconomics and risks	If effects on economics and social processes are relevant to the decision as well as those on human health and the environment, the effects should be assessed in an integrated manner. Beyond balancing costs of a management action against benefits, integration may require consideration of services of nature, values and preferences, and non-market mechanisms

1.2 *Roles of mathematical models for health and environmental risk assessment of chemicals*

Mathematical models that predict the transport and transformation of contaminants in environmental media have been developed both for scientific purposes and as applied tools for policy making, implementation, and management. In recent years, these models have been used to address the following diverse tasks (Mackay et al. 2001):

- Comparison of relative fates of different chemicals
- Identification of important fate processes
- Estimation of overall persistence and residence time
- Estimation of potential for long-range transport (LRT)
- Estimation of environmental concentrations and exposures
- Determination of bioaccumulation in organisms and food web
- Evaluation of likely recovery times of contaminated environments
- Checking the consistency of monitoring data
- Screening and prioritizing chemicals
- In general, as a decision support tool documenting the sources and nature o contamination and feasible remedial strategies

The models to estimate environmental contaminant transport and transformation are based on mass balance equations which account for the production, loss, and accumulation of the contaminant within a specified control volume. Transport phenomena and physical, chemical, and biological transformations are represented within the framework of this fundamental

concept. The mass balance principle is expressed mathematically as a time-dependent differential equation:

$$\frac{dM}{dt} = \frac{dCV}{dt} = M_{in}(C,t) - M_{out}(C,t) + S(C,t) \pm Rxn(C,t) \quad (1.1)$$

where M is the mass of the contaminant within the control volume, V is its volume, C (=M/V) is the concentration, t is time, M_{in} and M_{out} are the transport rates across the boundaries of the control volume from and to the surrounding environment, respectively, S is the source emission rate, and Rxn is the rate of internal reactions that may either produce or consume the
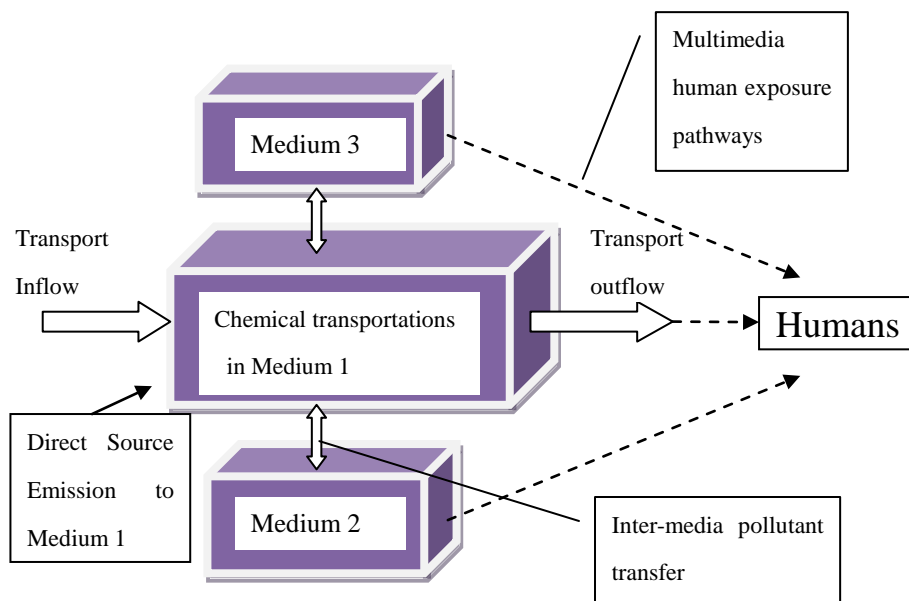
Figure 1.1: The schematic of mass-balance principle (based on Ramaswami et al. 2005)

contaminant. Figure 1.1 illustrates the mass-balance principle within a well-defined control volume.

In general contaminant transport and transformation models can be classified according to a number of contrasting characteristics (Ramaswami et al. 2005; Mackay et al. 2001):

- *Single-media versus Multimedia models*: To address media-specific problems, single-media models for air, surface water, groundwater, and soil pollution have been used by different disciplines. Although these models generally provide a more detailed description of the pollutant distribution in space and time and incorporate mass transfer from other media as boundary conditions, they are not capable of characterizing the total environmental impact of a pollutant release. Multimedia fate, transport, and

exposure models have been currently developed. These either link a number of single-media models for multiple exposure pathways within a single computing platform, or model the simultaneous partitioning of chemicals among multiple compartments.

- *Dynamic versus Steady state*: Dynamic models predict changes in environmental concentrations over time and consider temporal variations in source release and other model inputs. Steady-state models consider constant inputs and conditions where the reaction, transport, and mass-transfer terms are in equilibrium.

- *Eulerian (or box) versus Lagrangian models*: Eulerian models consist of a number of volumes or boxes, which are fixed a space and are usually treated as being homogeneous, i.e., well-mixed, in chemical composition. Lagrangian modes are to follow mathematically a pollution parcel in air or water as the parcel moves from place to place. These models consider the situation of heterogeneity in concentration and are suitable to set up diffusion/advection/reaction differential equations.

- *Analytical versus Numerical models*: Models that have analytical solutions generally require significant simplifying assumptions, e.g., steady-state conditions and spatial homogeneity. Models that need numerical solution consider significant spatial heterogeneity, temporal variability, or nonlinear transformation processes. These numerical techniques solve differential equations accounting for the chemical mass balance over discrete cells in space and/or discrete time steps.

- *Deterministic versus Stochastic models*: In deterministic models, a single concentration for each location and time in the model is calculated. Stochastic models produce a range or distribution of values for each prediction. This distribution may represent variability reflecting the temporal or spatial variation occurring in the environment, or uncertainty reflecting the imperfect knowledge on model inputs.

Currently multimedia mass balance models are increasingly being used to understand and evaluate the fate of chemicals in the environment, including the regulatory community (Webster et al. 2004b). In the context of the environmental health risk assessment, the multimedia models have been developed as essential tools, e.g., CalTOX (McKone 1993a), USES-LCA (Van Zelm et al. 2009), and 2-FUN tool (documents are available at www.2-fun.org). The integrated study of contaminant release, transport, fate, exposure, dose, and response is the basis for the environmental health risk assessment paradigm. The multimedia models can be used in line with the exposure and risk assessment paradigm, illustrated in Figure 1.2.

The first step in a health risk assessment is the estimation of ambient concentrations in each media (i.e., air, water, soil, vegetation, and etc) based on pollutant emissions, considering the fate and transformation processes of the pollutants. The
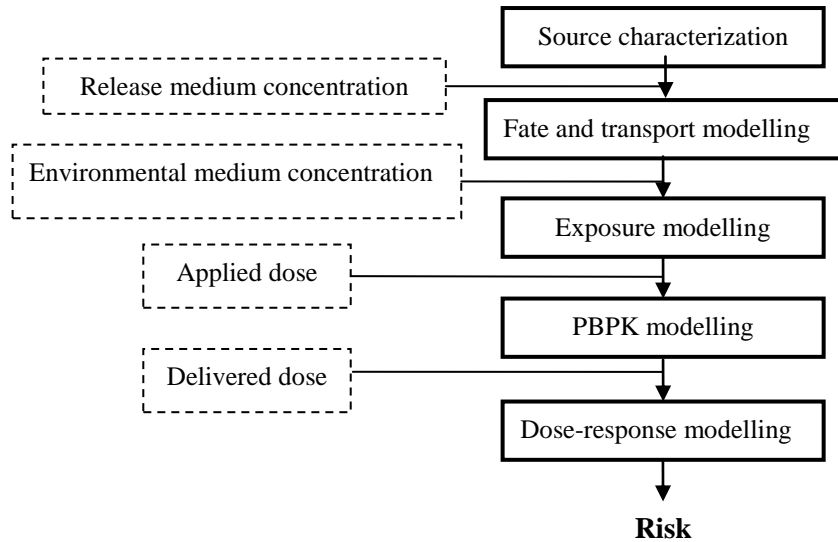
Figure 1.2: Integrated environmental health risk assessment scheme (based on Ramaswami et al. 2005)

calculated environmental concentrations are linked to exposure by using exposure models or exposure assessments. Principal pathways of exposure include inhalation of ambient air (both outdoor and indoor), ingestion of drinking water, contaminated foods, or contaminated soil, and dermal exposure to the air, soil, and domestic or recreational waters. The exposure for each pathway is calculated by the product of the pollutant concentration in the contact medium and an intake factor representing, for example, the breathing rate, water ingestion rate, food consumption rate, soil ingestion rate, or dermal contact rate. Some multimedia models, e.g., CalTOX, USES-LCA, and 2-FUN tool, have integrated systems to model both the environmental concentration of a pollutant in each medium and its exposure to humans via multi-pathways.

The actual effective dose received by a human body depends on how much of the pollutant is retained in the body and how much of it reaches target organs or tissues. There are some cases where the calculation involving the transformation from exposure to effective dose is made by simple statistical or conservative assumptions. For example, the modelling scheme of CalTOX and USES-LCA applies such simple statistical or assumptions. Using sophisticated physiologically based pharmacokinetic (PBPK) models allows the prediction of distribution and accumulation of contaminants in humans, considering their respiratory, digestive, and

circulatory systems to target organs and tissues. 2-FUN tool incorporates the PBPK models into its integrated system.

The final step of the health risk assessment is to link the effective dose to target organs and tissues with the health impact through dose-response functions. Functional relationships can range from a simple linear to more sophisticated relationships. These relationships can be estimated from past human exposures or high occupational doses, animal experiments, or epidemiological studies

1.3 *Consideration of parametric uncertainty/variability in health and environmental risk assessment*

Environmental systems contain highly variable properties. There is a great deal of uncertainty about these properties, inputs, and then responses. In definition, *variability* denotes inherent differences in environmental properties that exist over space and time (e.g., the differences in exposure, vulnerability, and risk that occur between one individual to another in a target population). On the other hand, *uncertainty* refers to a lack of knowledge about environmental processes and properties. Currently, many tools of probability and statistics are applied to characterize variability and uncertainty.

When modelling studies deal with health and environmental risk characterization of toxic chemicals, the consideration of parametric uncertainty/variability can be important and its importance has been well recognized (Solomon et al. 2000; Dubus et al. 2003; and Gutièrrez et al. 2009). A method commonly used to account for parametric uncertainty/variability is to characterize a targeted variable (parameter) by a Probability Density Function (PDF). In probability theory, the PDF of a continuous random variable describes the relative chance for this random variable to occur at a given point (value) in the observation space. The PDF must be integrated over a given range to find the probability of occurrence over that range:

$$\text{Prob}[x_1 < \theta < x_2] = \int_{x_1}^{x_2} fx(x)dx \qquad (1.2)$$

where θ denotes the name of the random variable (parameter, hereafter) and x is the particular value of θ. Common forms of the probabilistic density functions include the uniform, exponential, gamma, normal, and log-normal distribution.

In general, derivation of a PDF for a parameter requires a great deal of effort for the data collection. When the information about a given parameter is searched for, several situations can

come out, e.g., the situations where a large set of homogeneous data is available, where a large dataset of data is available but the data quality is heterogeneous (e.g., data obtained under different experimental protocols), where only a limited set of data is available, where no data is available for the specific substance and analogies and/or regression models must be used, and etc.

When a large set of data with homogeneous quality is available: For some chemical substances and their transfer factors in the environment, a large set of data can be obtained from the literature. In such cases, the collected abundant data can be directly fit to a PDF by classical statistical approaches.

When a large set of data with heterogeneous quality is available: Durrieu et al (2006) described a procedure to treat the 'quality' of data to derive a PDF. This procedure can be subdivided into the following stages:

i) Each of data is given the score that indicates the quality of data. Several criteria quantify the score given to each of data. The criteria include, for example, number of data replicates, physico-chemical characteristics measured in the experiment, and quality of the publication (e.g., a project report VS an article published in a peer-reviewed journal). On the basis of expert judgement integrating all the criteria, a database containing both collected data and their associated scores is built up.

ii) A direct weighted bootstrap procedure (detailed descriptions of the bootstrapping method can be found in Cullen and Frey (1999), Davison et Hinkley (1997), and Efron and Tibshirani (1993)) is then used to perform a weighted-sampling based on collected data. The probability of drawing data depends on their scores previously defined.

iii) The generated samples by the bootstrap procedure is then fitted to a theoretical distribution (e.g., Log-normal distribution). The goodness of fit can be tested by different criteria such as the Kolmogorov-Smirnov test with a Dallal-Wilkinson approach and the multiple r-square coefficient (r^2).

When only a limited set of data is available: Bayesian approaches can be of great interest to estimate a PDF when only a limited set of data is available for an investigated parameter. The Bayesian approaches require the prior knowledge of the parameter (e.g., the information obtained from previous studies) with actual data of the parameter (in this case, a poor dataset) and then calculate the posterior distribution (the posterior PDF). A couple of Bayesian approaches are described in detail in Chapter 3.

1.4 *Model review*

The current methodologies to assess the impacts on the human health and ecosystems of toxicants emitted to the environment are reviewed in this section. Main focus of this review is set at the approaches using mathematical models. However, non-modelling approaches are also described here to allow a more comprehensive review. Each of these approaches is summarized and evaluated by a fixed number of principal characteristics:

- **Impact categories (model outputs):** eco-toxicity impacts and/or human toxicity impact
- **Exposure routes:** ingestion, inhalation, dermal
- **Fate, exposure and effect:** if fate, exposure and effect analyses are included or not
- **Chemicals considered**: organic pollutants and/or metals
- **Media considered:** air, water (fresh, ground, sea...), soil, sediment, vegetation, food chain and etc
- **Spatial variation:** regional scale, continental scale, global scale, country and seas boundaries
- **Source code availability**
- **Model availability:** pay model or free model
- **Availability for sensitivity and uncertainty analyses**
- **Population category:** if the differences in man/woman and adult/child are considered or not

CalTOX

The CalTOX model was originally developed as a set of spreadsheet models and spreadsheet data sets for assessing human exposures from continuous releases to air, soil, and water (Mckone 1993a). Hertwich (Hertwich 1999; Hertwich et al. 2001) applied the CalTOX model for the assessment of human toxicity in LCA. Ecotoxicity is not evaluated in the model.

The current version of CalTOX (CalTOX4) is an eight-compartment regional and dynamic multimedia fugacity model. CalTOX comprises a multimedia transport and transformation model, multi-pathway exposure scenario models, and add-ins to quantify and evaluate variability and uncertainty. To conduct the sensitivity and uncertainty analyses, all input parameter values to CalTOX are given as distributions, described in terms of mean values and a coefficient of variation, instead of point estimates or plausible upper values.

1. Fate and exposure analysis

The multimedia transport and transformation model is a dynamic model that can be used to assess time-varying concentrations of contaminants that are placed in soil layers at a time-zero concentration or contaminants released continuously to air, soil, or water. This model is used for determining the distribution of a chemical in the environmental compartments

The exposure model encompasses 23 exposure pathways, which are used to estimate average daily doses (inhalation, ingestion of foods, and dermal contact) within human population linked geographically to a release region.

2. Effect analysis

The human toxicity potential (HTP), a calculated index that reflects the potential harm of a unit of chemical released into the environment, is a indicator to present the individual lifetime risk. This indicator is based on both the inherent toxicity of a compound and its potential dose. The CalTOX scheme can calculate cancer and non-cancer HTP values for air and surface water emissions of 330 compounds.

The overview of the partitioning among the liquid, solid and/or gas phases of individual compartments is presented in Figure 1.3.
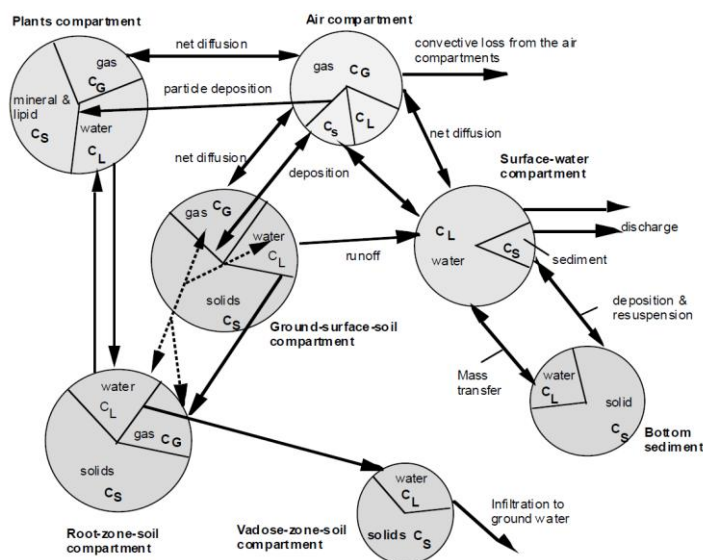
Figure 1.3: The overview of the partitioning among the liquid, solid and/or gas phases of individual compartments (Mckone 1993b). In the current version of CalTOX (CalTOX4), the plant compartment comprises two sub-compartments (plant surfaces (cuticle) and plant leaf biomass (leaves))

Table 1.2: Principal characteristics of Caltox model (based on Koning et al. 2002)

Principal characteristics	Multimedia model for fate analysis and extensive

15

	analysis of exposure pathways
Impact categories	Human toxicity
Exposure routes	Inhalation, ingestion, and dermal contact
Fate, exposure and effect	Fate, exposure, and effect are considered (the effect analysis is effect based)
Chemical considered	Organics and inorganics
Media considered	Air, water, sediments, 3 soil layers, vegetation
Spatial variation	Not considered
Source code availability	Yes, as Excel spreadsheet
Model availability	Yes
Dynamic or steady-state	dynamic
Availability for sensitivity and uncertainty analyses	Yes
Population category	Not considered

IMPACT 2002+

Based on the generalised framework established by the Society of Environmental Toxicology and Chemistry for Life Cycle assessment LCA (Jolliet et al. 1996b), IMPACT2002+ was developed to connect, as far as possible, each life cycle inventory (LCI) result (emissions or other intervention) to the corresponding environmental impacts (Jolliet et al. 2003a). The IMPACT 2002+ life cycle impact assessment methodology proposes a feasible implementation of a combined midpoint/damage approach, linking all types of LCI results via 14 midpoint categories to four damage categories (Figure 1.4). For IMPACT 2002+, new concepts and methods have been developed, especially for the comparative assessment of human toxicity and ecotoxicity.

Human toxicity

Human Damage Factors (HDF) are calculated for carcinogens and non-carcinogens, employing intake fractions, best estimates of dose-response slope factors, as well as severities. Indoor and outdoor air emissions can be compared and the intermittent character of rainfall is considered. The intake fraction accounts for a chemical's fate with respect to multimedia and spatial transport as well as human exposure associated with food production, water supply, and inhalation. This is then combined with an effect factor characterizing the potential risks linked to the toxic intakes. Severity characterizes the relative magnitude of the damage due to certain illness in the end. The Human Damage factor of substance i is then calculated as the production of intake fraction i and the corresponding exposure factor. The complete fate and exposure

assessment in IMPACT2002+ enables the estimation of a chemical's mass (or concentration) in the environmental media at a regional or at a global scale using the same basic model. Latest developments include the calculation of pesticide residues in food due to direct applications.

Ecotoxicity (aquatic and terrestrial)

Impacts on aquatic ecosystems are in many aspects considered in a similar manner with human toxicity concerning fate and effect. The difference is that fate enables to relate emissions to the change in concentration in the pure aqueous phase of freshwater and exposure is implicitly considered in the effect factor that represents the risk s at species level, finally leading to a preliminary indicator of damages on ecosystems. Terrestrial ecotoxicity potentials are calculated in a similar way with aquatic one.

Figure 1.5 shows the general scheme of the impact pathway for human toxicity and ecotoxicity.
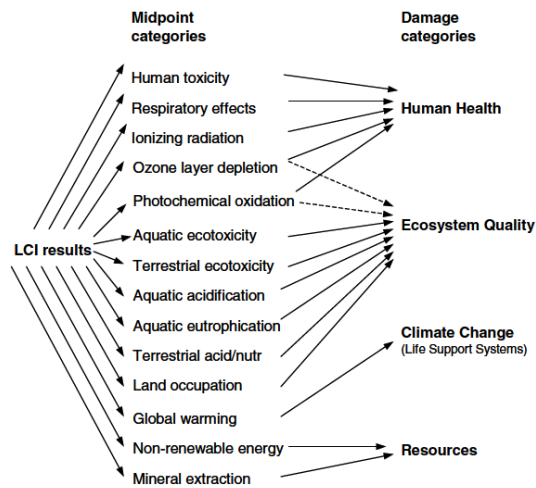
Figure 1.4: Overall scheme of the IMPACT 2002+, linking Life Cycle Impact (LCI) results via the midpoint categories to damage categories (Jolliet et al. 2003a)
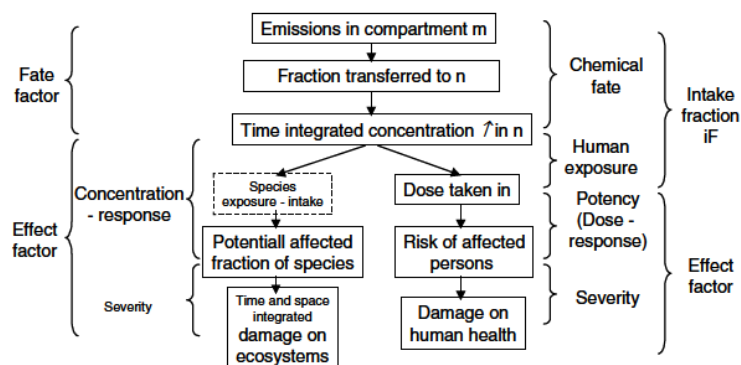
Figure 1.5: General scheme of the impact pathway for human toxicity and ecotoxicity (Jolliet et al. 2003b)

Table 1.3: Principal characteristics of IMPACT 2002+ model (based on Koning et al. 2002)

Principal characteristics	Multimedia chemical fate model combined with an exposure model for human health and potency/severity based effect analyses for human and ecotoxicological impacts
Impact categories	Human toxicity and ecotoxicity
Exposure routes	Inhalation and ingestion
Fate, exposure and effect	Fate, human exposure and toxicological effects are considered
Chemical considered	Predominantly for non-polar organics, but with adaptation to include metals, primary pollutants, particulate matter, and dissociating compounds
Media considered	Air, water (fresh and oceanic), soil, sediments, plants, and urban
Spatial variation	Regional and global scale
Source code availability	Logic and calculation procedure are fully documented
Model availability	Yes
Dynamic or steady-state	dynamic
Availability for sensitivity and uncertainty analyses	Yes
Population category	Not considered

USES-LCA

18

The model USES-LCA is a multimedia fate, exposure, and effects model. USES-LCA has been recently updated to USES-LCA 2.0 and contains a database of 3396 chemicals (Van Zelm et al. 2009). The model USES-LCA is based on the EUSES model family applied for risk assessment purposes in the European Union (Vermeire et al. 2005). This model is used to calculate characterization factors for ecotoxicity and human toxicity on both the midpoint and endpoint level. Characterization factors are used to determine the relative importance of a substance to toxicity related impact categories, such as human toxicity and freshwater ecotoxicity. The characterization factor accounts for the environmental persistence (fate) and accumulation in the human food chain (exposure), and toxicity (effect) of a chemical. In this method, for human toxicity, characterization factors for carcinogens, for non-carcinogens, and overall characterization factors are provided. Separate ecotoxicological characterization factors are provided for terrestrial, freshwater, and marine ecosystems. To obtain an overall ecotoxicological characterization factor on endpoint level they are further aggregated on the basis of species density of terrestrial, freshwater, and marine ecosystems separately.

For ten emission compartments, including urban air, rural air, freshwater, and agricultural soil, USES-LCA 2.0 calculates by default environmental fate and exposure factors in multiple compartments and human intake factors concerning air inhalation and oral ingestion via food and water intake by an infinite time horizon.

The nested multimedia fate model Simplebox 3.0 developed by Den Hollander et al. (2004) and included in EUSES 2.0 (EC 2004) is used for fate and exposure analysis of substances. The model structure of Simplebox has been slightly adapted to meet LCA-specific demands. For example, USES-LCA approach uses two geographic scales 'continental' and 'global', discarding 'local' and 'regional' scales (Huijbregts et al. 2005). Figure 1.6 presents schematic representation of Simplebox 3.0.

For human effect and damage analyses, USES-LCA 2.0 calculates Human toxicological Effect and Damage Factors (HEDFs) per chemical with information related to exposure route (inhalation and ingestion), and disease type (cancer and non-cancer). HEDFs on endpoint level express the change in damage to the total human population, expressed as disability adjusted life years (DALY), due to a change in steady-state exposure of the total human population. The HEDF consists of a disease-specific slope factor, and a chemical-specific toxic potency factor that reflects the average toxicity of a chemical towards humans.

For ecological effect analysis, Ecotoxicological effect factors (EEFs) on the endpoint level are used to express the change in overall toxic pressure due to a change in the concentration in a chemical. The EEF consists of a slope factor and a chemical-specific toxic potency factor that reflects the average toxicity of a chemical towards ecosystems.

Overview of USES-LCA 2.0 approach is presented in Figure 1.7.
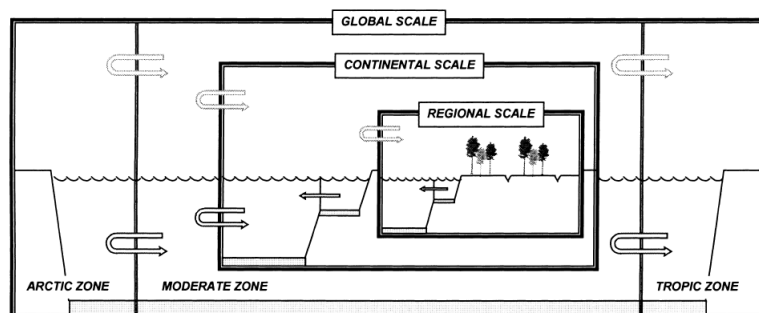
Figure 1.6: Schematic representation of Simplebox (Brandes et al. 1996). In USES-LCA 2.0, regional scale is not considered.
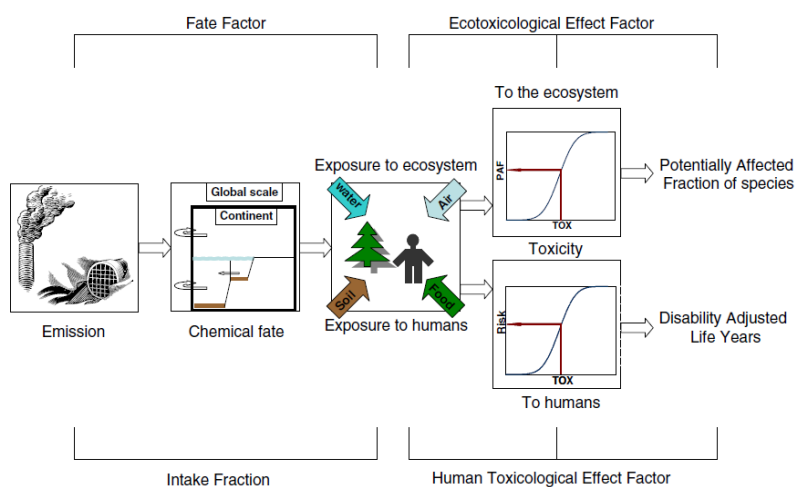
Figure 1.7: Overview of USES-LCA 2.0 approach (Van Zelm et al. 2009)

Table 1.4: Principal characteristics of USES-LCA methodology (based on Koning et al. 2002)

Principal characteristics	Multimedia model based on Simplebox 3.0
Impact categories	Human toxicity and ecotoxicity
Exposure routes	Inhalation and ingestion
Fate, exposure and effect	Fate, human exposure and toxicological effects are considered
Chemical considered	Organic and inorganic

Media considered	Global: air, (sea)water, and soil
	Continental: air, fresh water, seawater, natural soil, agricultural soil, industrial soil, fresh water sediment and marine water sediment
Spatial variation	Distinction between continental and global scale and between three climate zones (arctic, moderate, and tropic zones of the Northern hemisphere)
Source code availability	Yes, as Excel program
Model availability	Yes
Dynamic or steady-state	Steady-state
Availability for sensitivity and uncertainty analyses	Uncertainty analysis is available
Population category	Not considered

GLOBOX

As in the case of USES-LCA method, the GLOBOX is based on the EUSES 2.0 model (Sleeswijk and Heijungs 2010). It has primarily been constructed for the calculation of spatially differentiated LCA characterization factors on a global scale. Compare with the USES-LCA method (and EUSES model), the GLOBOX has a higher level of spatial differentiation in such a way that the GLOBOX is spatially differentiated with respect to fate and human intake on the level of separated, interconnected countries, and oceans/seas, whereas in the USES-LCA method, the evaluative region at the continental level (Western Europe) is not spatially differentiated. A main goal of the GLOBOX is to construct location-specific characterization factors for any emissions at any locations over the world, considering summed impacts of such emissions in different countries and seas/oceans.

The GLOBOX consists of the following three main modules:

- An impact-category independent fate module:

  Multimedia transport and degradation calculations are largely based on EUSES 2.0, and enhanced by site-specific (different countries/seas) equations to account for environmental advective air and water transport. Twelve distribution compartments are distinguished. Compared to EUSE 2.0, salt lakes, salt lake sediments, and groundwater are additional compartments.

- A human-intake module, applicable to all impact categories that are related to human intake of chemicals:

In this method, the values for the human body weight are taken into account to represent regional differences (e.g., difference between African countries and Vatican City). For determining the intake by humans, a consumption-based approach accounting for the intake per kilogram of body weight is considered, instead of a production (chemical)-based approach.

- An effect module, in which toxicity-related parameters can be introduced for every separate impact category:

In this method, two new, both chemical- and impact category-specific parameters are introduced; the sensitivity factor (SF) for ecotoxicity and the threshold factor (TF) for both ecotoxiciy and human toxicity. The SF reflects the fraction of area that is sensitive to a certain chemical, whereas the TF reflects the fraction of sensitive area where a predefined no-effect level for the ecosystem concerned, e.g. the hazardous concentration 5%.

In the end the GLOBOX combines fate-, intake and effect factors to calculate region-specific toxicity characterisation factors.

Table 1.5: Principal characteristics of Globox model (based on Sleeswijk and Heijungs 2010)

Principal characteristics	Multimedia model based on EUSES 2.0
Impact categories	Human toxicity and ecotoxicity
Exposure routes	Inhalation and ingestion
Fate, exposure and effect	Fate, human exposure and toxicological effects are considered
Chemical considered	Organic chemicals and metals
Media considered	Air, rivers, freshwater lakes, salt lakes, groundwater, sea water, freshwater lake sediment, salt lake sediment, sea sediment, natural soil, agricultural soil, urban soil
Spatial variation	Distinction between 239 different countries and 50 different seas (global scale)
Source code availability	For internal use only
Model availability	Yes
Dynamic or steady-state	Dynamic and steady state (Dynamic calculations for the estimation of toxicity potentials for different time hozizons)
Availability for sensitivity and uncertainty analyses	Not specified
Population category	Not considered

ECOSENSE

ECOSENSE is an integrated atmospheric dispersion and exposure assessment model which performs the Impact Pathway Approach developed in the framework of the ExternE project (Krewitt et al. 1995; Krewitt et al. 1998). ECOSENSE was developed to support the assessment of priority impacts resulting from the exposure to airborne pollutants, i.e., impacts on health, crops, building materials, forests and ecosystems. The current version of ECOSENSE 4.01 covers 14 pollutants, including the 'classical' pollutants SO_2, NO_x, particulates, CO and ozone, as well as some of the most important heavy metals and hydrocarbons, but does not include impacts from radioactive nuclides. Impacts of 'classical' pollutants are calculated on local (50km around the emission source), regional (Europe-wide) and (northern) hemispheric scale.

The ECOSENSE model uses a reference environment database which, on the 10×10 km^2 and 50×50 km^2 grids for the local and regional scales, respectively grid, provides data on population distribution and crop production (obtained from the EUROSTAT REGIO database), total agricultural area, building materials. Based on a comparison between the data relating to the receptors for the analysed territory and the pollutant distribution, the ECOSENSE model can calculate the exposition level of receptors and the resulting impact. Finally, ECOSENSE determines the total monetary damage by assigning the corresponding monetary value to each estimated impact.

ECOSENSE model makes fate analysis with the help of three air quality models. Each of them is used for different modelling scales, i.e., local and regional scales. Therefore, among various existing exposure pathways, ECOSENSE considers only inhalation.
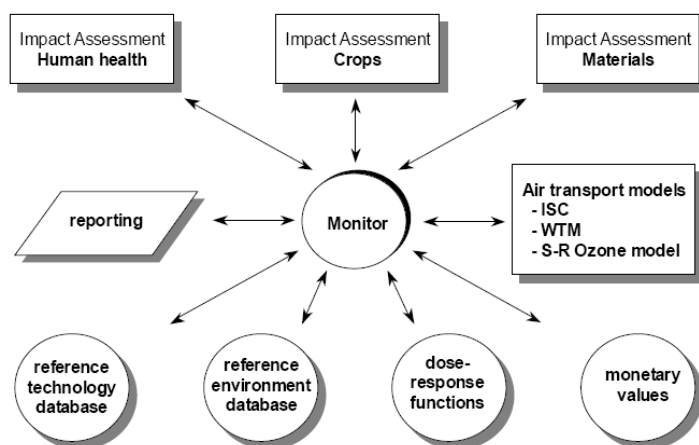
Figure 1.8: Modular structure of Ecosense model (Krewitt et al. 1995)

Table 1.6: Principal characteristics of Ecosense model (based on Koning et al. 2002)

Principal characteristics	Mixture of air transport models with impact analysis based on slopes of exposure-response relationship
Impact categories	Human toxicity
Exposure routes	Inhalation
Fate, exposure and effect	Fate (for air), human exposure and effect are considered
Chemical considered	Organic and inorganic chemicals
Media considered	Air, water, and soil
Spatial variation	Local and regional scales
Source code availability	No
Model availability	Yes
Dynamic or steady-state	Dynamic for air transport model
Availability for sensitivity and uncertainty analyses	Yes
Population category	Not considered

Ecopoints

The Ecopoints, or environmental scarcity, approach (Ahbe et al. 1990; Braunschweig et al. 1993) relates the concept of economic scarcity (relation between supply and demand) to scarcity of environmental absorption capacity. The scarcity of the environmental absorption capacity is given by the ratio between the actual anthropogenic emissions of a substance and the critical emission of the substance.

The equation to calculate the Ecofactor for a substance i comprises the normalization by the critical flow and the valuation using the linear relation between actual and critical as follows;

$$\text{Ecofactor}_i = \frac{1}{FK_i} \times \frac{F_i}{Fk_i} \times c \qquad (1.3)$$

Fk_i: the critical flow of substance i in a country, in a time period;

F_i: the actual flow of substance i, in a country, in a time period;

c: a dimensionless factor (10^{12}) to avoid large negative exponent values

F and Fk are given in tons, kWh, m^3 or other physical units per year. The unit of the Ecofactor is expressed as Ecopoints per physical unit of emitted substance. The critical flow (Fk) represents the absorption capacity of an environmental compartment for a particular substance. The actual flow (F) is the current total emission, independent of the investigated process or product.

For determination of the critical flow, the method uses politically defined and scientifically supported standards as bases of valuation. Those standards may be set nationally (clean air policy, clean water policy, acidification policy, eutrophication policy, and etc.) or internationally (international treaties and protocols).

The environmental index I (Ecopoints) is simply calculated as multiplying all emissions (which occur during the life cycle of the product P under investigation) by their corresponding Ecofactors and then adding them up:

$$I = \sum (Ecofactor_i \times E_i) \qquad (1.4)$$

Where E_i represents the quantity of the substance I emitted.

There are no fate and exposure analyses in the Endpoints method. The method has been elaborated for 10-30 different emissions into air and water.

Table 1.7: Principal characteristics of Ecopoints methodology (based on Koning et al. 2002)

Principal characteristics	No modelling; distance-to-target method using policy standards
Impact categories	No separate categories distinguished; only Ecopoints
Exposure routes	Not considered
Fate, exposure and effect	Only effect through policy standards
Chemical considered	Organic and inorganic chemicals
Media considered	Not considered
Spatial variation	Different sets of ecofactors for different countries
Source code availability	Not specified
Model availability	Not specified
Dynamic or steady-state	Not considered
Availability for sensitivity and uncertainty analyses	Not considered
Population category	Not considered

Eco-indicador 99

The Eco-indicador 99 methodology has been developed by Hofstetter (1998) and Goedloop and Spriensma (1999). The most critical and controversial step in the Life Cycle Impact Assessment (LCIA) is the weighting step. With this in mind the Eco-indicator methodology has been developed top-down regarding the weighting step as a starting point. Two important requirements for the weighting step need to be considered, if the weighing has to be conducted by a panel:

1. The number of impact categories (environmental damages) to be weighted should be as small as possible.
2. The impact categories should be concrete and easy to be explained to a panel.

Three environmental damages (end-points) to be weighted were selected based on two requirements above:

- Damage to human health
- Damage to ecosystem quality
- Damage to resources

The following damage models have been established to link these impact categories with the inventory result that is presented by resources, land-use, and emission of chemical substances.

Damages to human health are expressed as the Disability Adjusted Life Years (DALY). Models have been developed for respiratory and carcinogenic effects, the effects of climate change, ozone layer depletion, and ionising radiation. There are four sub steps in the models for human health:

  i) Fate analysis; linking an emission (expressed as mass) to a temporary change in concentration
  ii) Exposure analysis; linking the temporary change to a dose
  iii) Effect analysis; linking the dose to a number of health effects, i.e., the number and types of cancers
  iv) Damage analysis; linking health effects to DALYs

Damages to ecosystem quality are represented by two different expresssions:

  - For ecotoxicity, the Potentially Affected Fraction (PAF) is used. The PAF that can be interpreted as the fraction of species that is exposed to a concentration equal to or higher

than the No Observed Effect Concentration (NOEC). PAFs are based on substance-specific species-sensitivity distributions.

- For acidification, eutrophication, and land-use/transformation, the Potentially Disappeared Fraction (PDF) is used. The PDF indicates the effects on vascular plant populations in an area, i.e., the fraction of species that has a high probability of disappearance in a region due to unfavourable conditions.

For ecotoxicity, acidification, and eutrophication, the following sub steps are considered in the similar manner as for human health:

i) Fate analysis; linking emissions to concentrations

ii) Effect analysis; linking concentrations to toxic stress (PAF) or increased nutrient/acidity levels

iii) Damage analysis; linking the effects to the increased potentially disappeared fraction for plants (PDF)

Land-use/transformation is modelled on the basis of empirical data of the effect on vascular plants, as function of the land-use type and the area size.

Damages to resources is modelled in two steps

i) Resource analysis; which can be regarded as a similar step as the fate analysis, as it links an extraction of a resource to a decrease of the resource concentration

ii) Damage analysis; linking the decreased resource concentration to the increased effort to extract the resource in the future
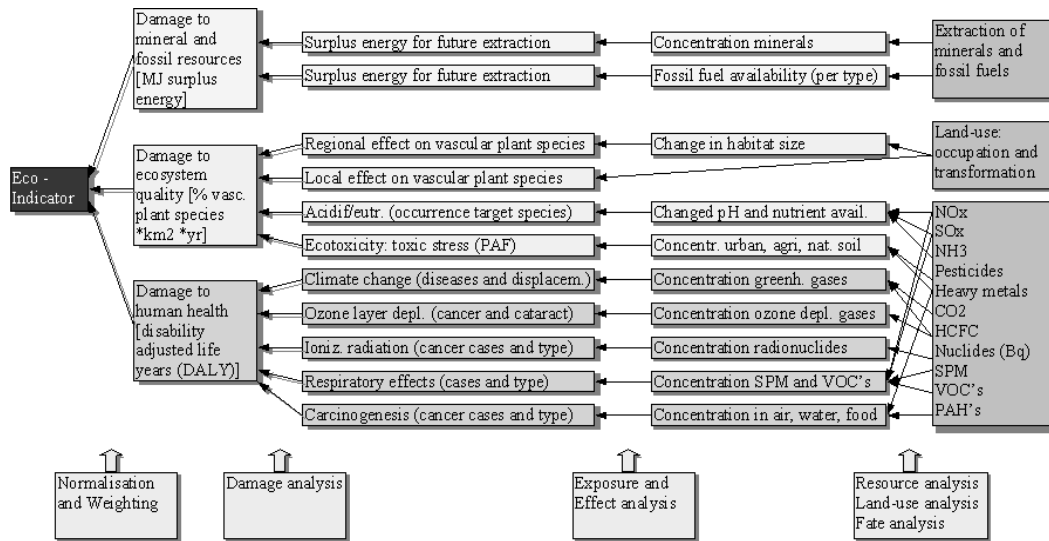
Figure 1.9: General representation of the Eco-indicator 99

Table 1.8: Principal characteristics of ECO indicator 99 model (based on Koning et al. 2002)

Principal characteristics	Multimedia model based on EUSES using the DALY and PAF concepts
Impact categories	Human toxicity and ecotoxicity
Exposure routes	Inhalation and ingestion
Fate, exposure and effect	All are considered + damage analysis
Chemical considered	Organic and inorganic chemicals
Media considered	Air, water, natural soil, agricultural soil, and industrial soil
Spatial variation	Regional scale
Source code availability	Not specified
Model availability	Yes
Dynamic or steady-state	Steady-state
Availability for sensitivity and uncertainty analyses	Uncertainly analysis is available
Population category	Not considered

EDIP-Characterisation

EDIP97 is a thoroughly documented midpoint approach covering most of the emission-related impacts, resource use and working environment impacts (Hauschild and Wenzel 1998) with normalization based on person equivalents and weighting based on political reduction targets for environmental impacts and working environment impacts, and supply horizon for resources.

Without using integrated, quantitative models, ecotoxicity and human toxicity are estimated using a simple key-property approach where the most important fate characteristics such as inter-media transport and atmospheric degradation/biodegradation, intake, and effect are included in a simple modular framework requiring relatively few substance data for calculation of characterization factors.

The update version of EDIP methodology (EDIP2003) (Hauschild and Potting 2005) supports three levels of spatial differentiation in characterization modelling; site-generic (no spatial differentiation), site-dependent (the level of countries or regions within countries), and site-specific (local scale) modelling. Apart from the feature of spatial differentiation, the EDIP2003 characterization factors take a larger part of the causality chain into account for all the non-global impact categories. Thus it includes the modelling of the dispersion and distribution of the substance, the exposure of the target systems to allow assessment of the exceedance of thresholds.

Table 1.9: Principal characteristics of EDIP-Characterisation methodology (based on Koning et al. 2002)

Principal characteristics	Integrated quantitative models focusing on independent environmental key properties
Impact categories	Human toxicity and ecotoxicity
Exposure routes	Inhalation and ingestion
Fate, exposure and effect	All but only partially
Chemical considered	Organic and inorganic chemicals
Media considered	Air, water, natural soil, agricultural soil, industrial soil
Spatial variation	Not considered
Source code availability	No
Model availability	Not specified
Dynamic or steady-state	Not specified
Availability for sensitivity and uncertainty analyses	Not considered
Population category	Not considered

USETOX

In 2005, a comprehensive comparison of life cycle impact assessment toxicity characterisation models was initiated by the United Nations Environment Program (UNEP)–Society for

Environmental Toxicology and Chemistry (SETAC) Life Cycle Initiative. The main objectives of this effort were (Rosembaum et al. 2008):

- To identify specific sources of differences between the models' results and structure
- To detect the indispensable model components
- To build a scientific consensus model from them, representing recommended practice

This led to the development of USETOX, a scientific consensus model that contains only the most influential model elements. These were, for example, process formulations accounting for intermittent rain, defining a closed or open system environment or nesting an urban box in a continental box.

Table 1.10: Models compared to build USETOX consensus models

Model	Strong point(s)	Weak point(s)
CalTOX	Most encompassing in terms of exposure pathways Advanced modelling of soil (several layers) Monte Carlo uncertainty estimation	No severity measure for human toxicity, only partly compatible with damage approach Ecosystem toxicity not assessed (e.g., marine environment and coastal zone not included for fate modelling)
IMPACT 2002	Continental average characterization factors available for different global regions Considering indoor air exposure Direct application of pesticides considered HC50 approach for effect modelling	Marine environment poorly represented so far
USES-LCA	Marine environment included Global coverage, however, not spatially resolved HC50 approach for effect modelling One-dimensional uncertainty factors available	
BETR	Flexible structure allows for spatially explicit chemical fate assessment at a variety of scales.	Chemical fate model only. No integrated ecotoxicity assessment. No multi-pathway human exposure assessment
EDIP	Key property based	Mainly representative for Europe

	Normalization and weighting methods provided	No explicit fate results available No severity measure for human toxicity
WATSON	European-wide spatially resolved fate, exposure and impact assessment (bottom-up analysis) Monetary valuation as weighting method	Confined to Europe Open system boundaries Ecosystem toxicity not assessed At present, confined to non-volatile compounds
EcoSense	Most reliable modelling of classical air pollutants amongst the chosen models Bottom-up, i.e., spatially resolved, assessment capabilities for Europe, Russia, China/Asia, and Brazil/South America Monetary valuation as weighting method	Open system boundaries Organic chemicals mostly not considered Only inhalation exposures with respect to toxic impacts (additionally impacts on crops and building materials) Ecosystem toxicity not assessed

USETOX was developed following a set of principles including (Rosenbaum et al. 2008):

- Parsimony: as simple as possible, as complex as necessary
- Mimetic: not differing more from the original models than these differ among themselves
- Evaluated: providing a repository of knowledge through evaluation against a broad set of existing models
- Transparent: being well-documented, including the reasoning for model choices

USETOX calculates characterisation factors for human toxicity and freshwater ecotoxicity. Assessing the toxicological effects of a Chemical emitted into the environment implies a cause–effect chain that links emissions to impacts through three steps: environmental fate, exposure and effects. Linking these steps, a systematic framework for toxic impacts modelling based on matrix algebra was developed within the OMNITOX project (Rosenbaum et al. 2007). USETOX covers two spatial scales (Figure 1.10).
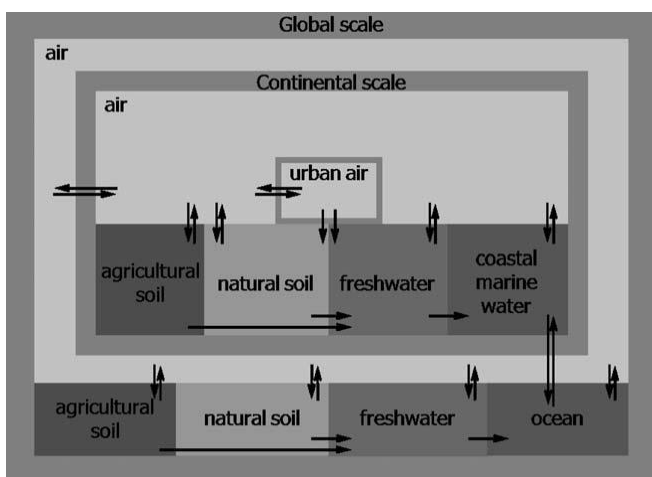
Figure 1.10: Compartments setup of USETOX (Rosembaum et al. 2008)

USETOX provides a parsimonious and transparent tool for human health and ecosystem characterization factor (CF) estimates. It has been carefully constructed as well as evaluated via comparison with other models and falls within the range of their results whilst being less complex.

Table 1.11: Principal characteristics of USETOX

Principal characteristics	A scientific consensus model based on comparison of seven models
Impact categories	Human toxicity and freshwater ecotoxicity
Exposure routes	Inhalation and ingestion
Fate, exposure and effect	All considered
Chemical considered	Organic chemicals
Media considered	Continental scale: Urban air, rural air, agricultural soil, industrial soil, freshwater, and coastal marine water Global scale: same structure as the continental scale (without urban air)
Spatial variation	Continental and global scales
Source code availability	Not specified
Model availability	Yes
Dynamic or steady-state	An algorithm estimating the effect of intermittent rain events is included to overcome the overestimation by steady-state scheme
Availability for sensitivity and uncertainty analyses	Not considered

Population category	Not considered

WMPT

The Waste Minimization Prioritization Tool (WMPT) is a computerized system developed by the U.S. EPA to assist in the ranking of chemicals for pollution prevention efforts in terms of their persistence, bioaccumulation and toxicity (PBT) properties (USEPA 1997b). WMPT contains screening level information on several thousand chemicals and can be used for a variety of prioritization applications.

The WMPT scoring algorithm provides chemical-specific scores that can be used for screening-level risk-based ranking of chemicals (the results from the algorithm should not be treated as a substitute for a detailed risk assessment). Figure 1.11 illustrates that the scoring algorithm is designed to generate an overall chemical score that reflects a chemical's potential to pose risk to either human health or ecosystems. A measure of human health concern is derived by jointly assessing the chemical's human toxicity and potential for exposure. In a similar manner, a measure of the ecological concern is derived from jointly assessing the chemical's ecological toxicity and potential for exposure.

As shown in Figure 1.11, scores are first generated at sub-levels. These scores at sub-levels are then aggregated upward to generate an overall chemical score. A score for a given "subfactor" (i.e. cancer effects) is derived by evaluating certain "data elements" that appropriately represent the subfactor. For example, one of the subfactors, "persistence", is scored by using a steady-state, non-equilibrium multimedia partitioning model to estimate regional half-life.
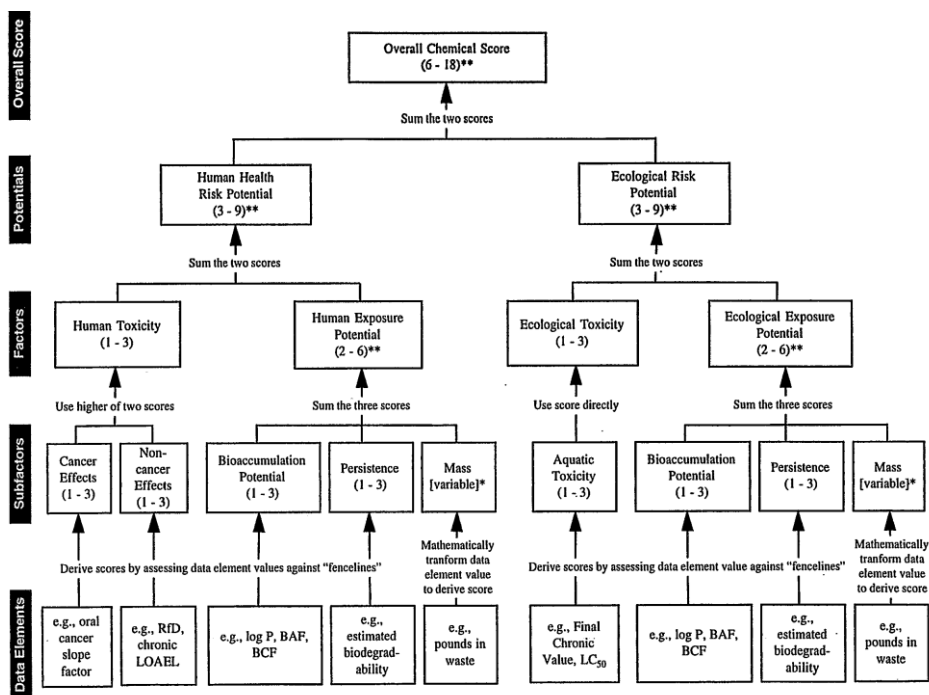
Figure 1.11: Overview of the WMPT scoring (USEPA 1997b)

Table 1.12: Principal characteristics of WMPT

Principal characteristics	Screen-level/ risk-based ranking tool
Impact categories	Human toxicity and ecotoxicity
Exposure routes	Not specified
Fate, exposure and effect	The method is performed by summing scores
Chemical considered	Organic compounds and inorganic compounds
Media considered	air, water, soil, sediment
Spatial variation	Not considered
Source code availability	No
Model availability	Yes
Dynamic or steady-state	Steady-state
Availability for sensitivity and uncertainty analyses	No
Population category	Not considered

ChemCAN model

The ChemCAN model describes the fate of a chemical in a region, assuming steady state conditions in the environment. The model estimates average concentrations in four primary environmental media consisting of air, surface water, soil, and bottom sediment, and three

secondary media consisting of groundwater, coastal water, and terrestrial plants. Chemical fate is determined through the solution of the set of mass balance equations for the primary media as described by Mackay (1991). The model is intended to assist in human exposure assessment where a specific target population may be identified.

Originally designed for use in Canada, a database of 24 regions of Canada is available. Other regions can be defined by the user. In the model, the appropriate dimension of surface areas is set between 100×100 km^2 and 1000×1000 km^2. The regional divisions of Canada were based on the eco-zones identified by Environment Canada and with consideration of the distribution of population and industrial activity, political boundaries, drainage basins, and climate to give areas of sufficiently homogeneous ecological conditions such that meaningful assessments of chemical fate can be conducted.

The transparency of this model was achieved by making it possible for the user to view the equations within the model. By viewing the section of program code, the user can know how this steady-state model mimics the physical reality. The model is intended to provide regionally-specific estimates of chemical concentrations in the primary media. These estimates can be compared to monitoring data and used for exposure estimation.

A current application of this model was presented in Webster et al. (2004b)

Table 1.13: Principal characteristics of ChemCAN

Principal characteristics	Multimedia model with steady-state condition
Impact categories	Human toxicity
Exposure routes	Not specified
Fate, exposure and effect	Fate
Chemical considered	Organic compounds and involatile compounds
Media considered	air, surface water, soil, bottom sediment, groundwater, coastal water, and terrestrial plants
Spatial variation	Regional
Source code availability	Yes
Model availability	Yes
Dynamic or steady-state	Steady-state
Availability for sensitivity and uncertainty analyses	No
Population category	No

XtraFOOD model

The XtraFood model was developed in the research project initiated by VITO (Seuntjens et al. 2006). The model calculates transfer of contaminants in the primary food chain (Figure 1.12). In the project, the transfer model was coupled with historical food consumption data to estimate human exposure to contaminated food products. The model focuses on the terrestrial food chain. The XtraFOOD model consists of three modules, which are inter-linked:

- A mass balance model at the farm level: calculation of inputs and outputs;
- Bio-transfer module: calculation of the transfer of contaminants to vegetable products (vegetables, cereals, animal feed) and animal products (meat, milk dairy products, poultry, eggs);
- Exposure and impact module: calculation of the exposure from food (and other exposure routes) and comparison with reference values.

The XtraFOOD model calculates as output the food intake and resulting contaminant intake, independently for age and gender categories. Exposure can be calculated as being representative for a population or separately for local and background intake. All these intakes are linked to the model output. Additional intakes are provided to add concentration data in non-farm related foods (e.g., fruit juice, fish, and so on). Eventually human exposure can be calculated and compared to the available toxicological levels (thresholds) to estimate impacts on human health.
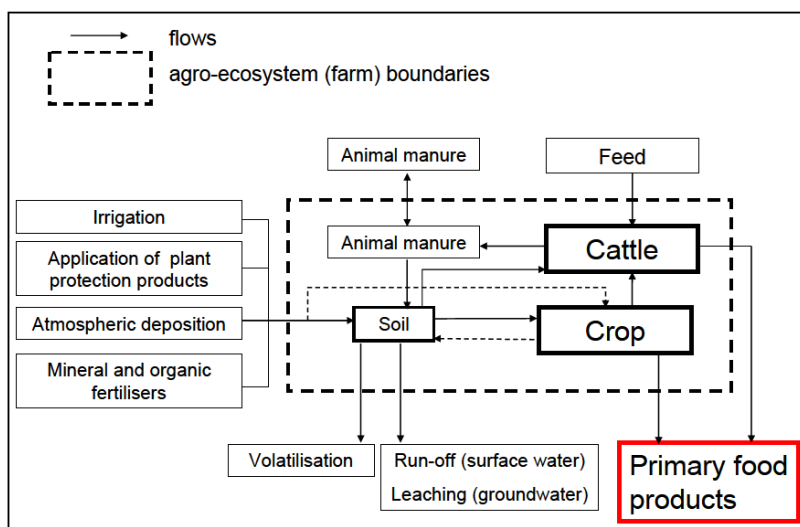

Figure 1.12: Overview of contaminant flows in a model agro-ecosystem to the food chain (Seuntjens et al. 2006)

Table 1.14: Principal characteristics of XtraFOOD

Principal characteristics	Multimedia model focused on the primary food chain
Impact categories	Human toxicity
Exposure routes	Ingestion
Fate, exposure and effect	Fate and exposure are considered
Chemical considered	Organic and heavy metals
Media considered	Air, soil, farm-related crops, animal
Spatial variation	Not considered
Source code availability	Not considered
Model availability	Not specified
Dynamic or steady-state	Steady-state
Availability for sensitivity and uncertainty analyses	Yes
Population category	Age and gender are considered

2-FUN tool

2-FUN tool is new integrated software based on an environmental multimedia model, physiologically based pharmacokinetic (PBPK) models, and associated databases. The tool is a dynamic integrated model and is capable of assessing the human exposure to chemical substances via multiple exposure pathways and the potential health risks (Figure 1.13). 2-FUN tool has been developed in the framework of the European project called 2-FUN (Full-chain and UNcertainty Approaches for Assessing Health Risks in FUture ENvironmental Scenarios: information is available at www.2-fun.org).

The environmental multimedia model contained in 2-FUN tool was developed based on the extensive comparison and evaluation of existing multimedia models such as CALTOX, SimpleBox, XtraFood, and etc. The multimedia model comprises several environmental modules, i.e., air, fresh water, soil/ground water, several crops, and animal (cow and milk). It is used to simulate chemical distribution in the environmental modules, taking into account the manifold links between them. The PBPK models were developed to simulate the body burden of toxic chemicals throughout the entire human lifespan, integrating the evolution of the physiology and anatomy from childhood to advanced aged. These models are based on a detailed description of the body anatomy and include a substantial number of tissue compartments to enable detailed analysis of toxicokinetics for diverse chemicals that induce

multiple effects in different target tissues. The key input parameters used in both models were given in the form of probability density function (PDF) to allow the exhaustive probabilistic analysis and sensitivity analysis in terms of simulation outcomes.

The environmental multimedia and PBPK models were built and linked together on the common platform software called Ecolego®(www.facilia.se). One of the main characteristic of Ecolego system is the use of 'Interaction matrix' (described in the section 2.1) to build and visualize models (Figure 1.14). The effective graphical simulation interface presented in the Ecolego system can facilitate a comprehensive identification and visualization of the exposure pathways and allow classification of the role of different environmental modules (subsystem) in terms of transfer relationship. In the Ecolego system, advanced methods concerning probabilistic and sensitivity analyses can be selected: (i) Monte Carlo methods for the propagation of parametric uncertainties, (ii) an optimization function to correlate input parameters with simulated outputs in the Monte Carlo process and then to optimize the values of input parameters, and (iii) several regression and Fourier approaches for conducting sensitivity analysis.

The complete 2-FUN tool is capable of realistic and detailed lifetime risk assessments for different population groups (general population, children at different ages, pregnant women), considering human exposure via multiple pathways such as drinking water, inhaled air, ingested vegetables, meat, fish, milk, and etc. In the scheme of an integrated environmental health risk assessment presented in Figure 1.2 (Section 1.2), 2-FUN tool can cover the scheme from 1st to 4th steps.
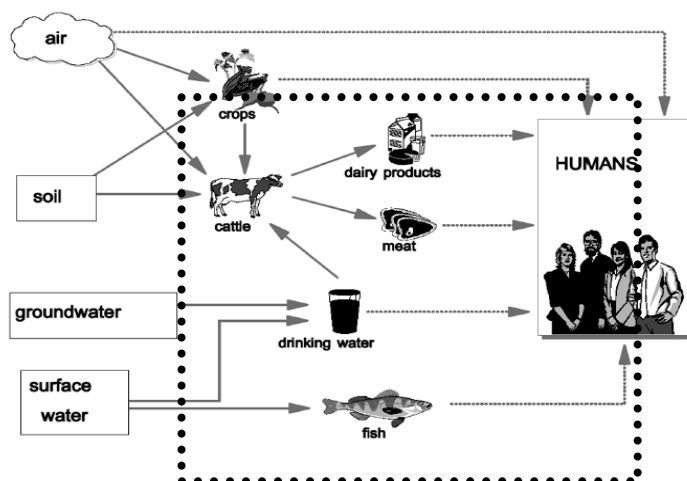
Figure 1.13: Multi-pathways that the substances can take into the human (The part enclosed by a dashed line emphasizes the indirect pathways into the human via food chains)
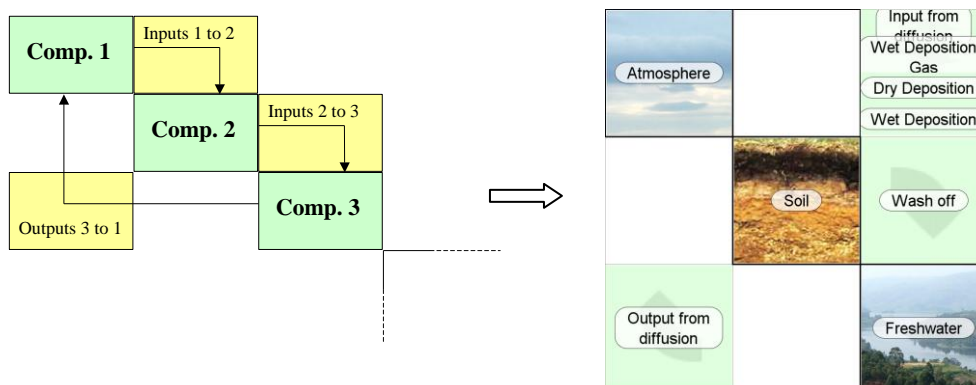
Figure 1.14: The schematic of interaction matrix (on the left) and the representation of interaction matrix in the Ecolego system (on the right)

Table 1.15: Principal characteristics of 2-FUN tool

Principal characteristics	Integrated tool coupling an environmental multimedia model and PBPK models
Impact categories	Human toxicity
Exposure routes	Ingestion, Inhalation, and dermal intake (planned)
Fate, exposure and effect	Fate, exposure, and potential effect are considered
Chemical considered	Organic and inorganic
Media considered	Air, fresh water, soil/ground water, farm-related crops, and animal (cow and milk)
Spatial variation	Not considered (mainly used for regional scale)
Source code availability	Yes
Model availability	Yes in the near future
Dynamic or steady-state	Dynamic
Availability for sensitivity and uncertainty analyses	Yes
Population category	Age and gender are considered

2-FUN tool has the following prominent features, which differentiate it from other models:

- Its capability to conduct a full-chain risk assessment on a common system, which allows linking the simulation of chemical fate in the environmental media, multiple pathways of exposure, and the detailed analysis for multiple effects in different target tissues in human body (by PBPK models)
- Its capability to assess the health risk of specific human groups vulnerable to toxicants, i.e., for woman, infant

39

- It contains a wide range of methods for sensitivity and uncertainly analyses.
- It contains an exhaustive database of PDF for input parameters.
- It can be user-friendly because of its effective graphical simulation interface and its flexibility, which facilitates users to design scenarios for target regions and arrange the tool on their own ways, i.e., users can select only the environmental modules necessary for their regional scenarios.

Selection of the models to perform risk assessments of human health and ecosystems should depend on the purposes and scenarios designed by model users. For example, the selection will be different depending on if the models are used for screen-level or detailed-level risk assessment, and if the scenarios are designed with large scale (global or continental) or small scale (regional or local).

1.5 *General objective*

As described in the former sections, modelling tools have been widely used for the environmental and health risk assessment, and the importance to consider the parametric uncertainty for environmental and health risk assessment has been well recognized. This dissertation then addresses three subjects. The first is the introduction of 2-FUN tool, where conceptual models defined in 2-FUN tool are described together with their mass balance equations (Chapter 2). The second is the investigation of statistical approaches to derive probabilistic density functions (PDFs) of the input parameters that are defined in 2-FUN tool (freshwater compartment), where Bayesian methods are mainly discussed (Chapter 3). The third is the application of 2-FUN tool for a designed case study in regional scale, where the chemical exposure to humans via multiple pathways is considered and the uncertainty and sensitivity analyses for model outputs (internal concentrations in organs in our case study) are discussed (Chapter 4).

2 STRUCTURE OF 2-FUN TOOL

2.1 *Conceptual description of 2-FUN tool*

As described in the section 1.4, 2-FUN tool consists of an environmental multimedia model, PBPK models, and associated databases. The first step in the development of the 2-FUN tool was the construction of a conceptual model defining the biosphere compartments, e.g., air, water, soil, crops, animals etc, which were eventually sub-divided in several sub-compartments,

and the relations between these compartments (i.e. transfers governed by physical, chemical and/or biological processes). The environmental multimedia model included in 2-FUN tool comprises five environmental compartments and a PBPK model corresponds to a human compartment:

- Air compartment
- Freshwater compartment
- Soil / groundwater compartment
- Plant compartments (representing edible plants compartments such as root, potato, leaf, grain and fruit)
- Animal compartment (representing meat and milk cows)
- Human compartment (PBPK model)

The 2-FUN tool considers the mass balance principle in the whole system as well as in each compartment. Interaction matrix is an expert qualitative method to identify multiple interactions among the multiple compartments, which facilitates the comprehensible identification and visualization of exposure pathways and allows the classification of roles of compartments in terms of transfer relationships.

The interaction matrix is a table which describes the conceptual model by tabulating the interactions between the compartmental media. The main compartments of the biosphere system are identified and listed in the leading diagonal elements (LDEs, green color) of the table, and the interactions between the LDEs are listed in the off-diagonal elements (ODEs, yellow color) (Figure 2.1).
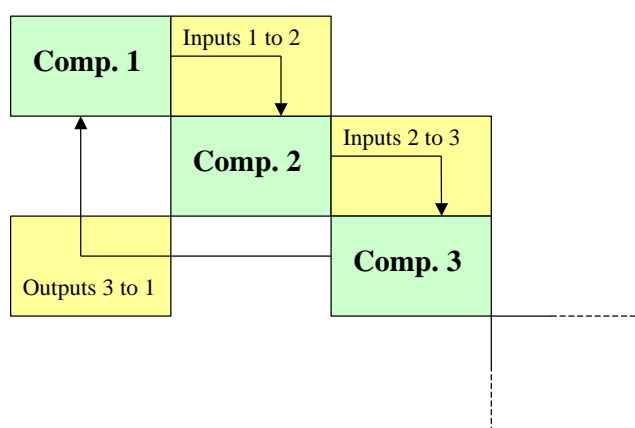

Figure 2.1: Schematic of interaction matrix

The 2-FUN conceptual model was defined independently for each constituent compartment, by reviewing existing models, frameworks and methodologies currently used for assessing

transfer of chemicals. The review a priori selected two types of models, frameworks, and methodologies:

- Integrated multimedia models, covering all the water, soil, and biota compartments and aiming at calculating generic human exposures
- Fate models, providing a detailed description of the behaviour of chemicals in a specific compartment

For example, the conceptual model of freshwater compartment was built up based on the following existing models:

- AQUATOX (USEPA 2004): Ecological food-web freshwater model kinetically describing transfer of chemicals in various abiotic and biotic compartments
- OURSON (Ciffroy 2006): Dynamic transfer initially developed for simulating the human exposure to radionuclides and metals discharged in freshwater
- OWASI (Mackay 1983; Di Guardo 2006; Warren et al 2007): Model simulating the steady-state chemical concentration in a lake or river segment
- SimpleBox: Steady-state multimedia model incorporated in the EUSES system
- TRIMFate (USEPA 2002): Compartmental mass balance model providing exposure estimates for ecological receptors (plants and animals), in particular in freshwater systems

In order to cover different approaches for the pollutant transfer in a given compartment, some contrasting tools were selected for the review (e.g., from screening models to detailed mechanistic ecological models). The analysis over the selected models raised some key issues in model mechanism and then permitted building an appropriate conceptual model by addressing those key issues. Figure 2.2 presents the 2-FUN interaction matrix for the freshwater compartment.

The detailed descriptions of the model review and development for all the conceptual models (i.e., air, freshwater, soil/groundwater, plants, and animal) are present in the deliverables of 2-FUN project (available at www.2-fun.org).

Atmosphere Gas			Dry and wet deposition					
	Atmosphere Aerosols	Dry and wet deposition	Dry and wet deposition					
		Soil surface	Erosion-runoff					
			River/lake water Dissolved phase	Adsorption		Bioaccumulation	Bioaccumulation	Physico-chemical and biological degradation
			Desorption	River/lake water Suspended particulate Matter	Particles deposition Water diffusion			
				Particles resuspension Pore water diffusion	Bottom sediments			
						Herbivorous fish	Biomagnification	Biological elimination Degradation
							Carnivorous fish	Biological elimination Degradation
								Sink

Figure 2.2: The 2-FUN interaction matrix for the freshwater compartment

2.2 *Mass balance equations*

As a next step, the conceptual models were translated into mathematical expressions. Time-dependent pollutant concentrations of different phases in each biosphere compartment (i.e. dissolved phase, suspended particulate matter, bottom sediments, and fish for freshwater compartment) were derived from the inputs, outputs, and transformation processes defined in conceptual models and mathematically described by mass balance systems. Only principal mass balance equations are presented in this section. All the detailed mathematical expressions for environmental compartments are described in the deliverables of 2-FUN project. Those for human compartment are described in detail in Beaudouin et al. (2010).

Freshwater compartment: It consists of two phases, i.e., raw river water and bottom sediments. The mass balance equation of a chemical in the water phase is given as follows:

$$\frac{dM_{rw}}{dt} = \underbrace{\frac{D_{point-sources}(t)}{86400 \cdot Q(t)}}_{\text{Point sources}} + \underbrace{Q(t) \cdot C_{rw_{upstream}}}_{\text{Input from upstream}} - \underbrace{Q(t) \cdot C_{rw}(t)}_{\text{Output to downstream}} + \underbrace{\frac{dM_{rw,dry_{dep}}}{dt}}_{\text{Input from dry particulate deposition}} +$$

$$\underbrace{\frac{dM_{rw,wet_{dep,part}}}{dt}}_{\text{Input from wet particulate deposition}} + \underbrace{\frac{dM_{rw,wet_{dep,gas}}}{dt}}_{\text{Input from gaseous deposition}} + \underbrace{\frac{dM_{air,diff_{gas}}}{dt}}_{\text{Input by diffusion}} -$$

$$\underbrace{\frac{dM_{rw,diff_{gas}}}{dt}}_{\text{Output by diffusion}} \underbrace{\frac{dM_{rw,wash-off}}{dt}}_{\text{Input from soil wash-off}} \pm \underbrace{\frac{dM_{rw,sed_{resus,dep}}}{dt}}_{\text{Input and output by sediment resuspension and deposition}} \pm$$

$$\underbrace{\frac{dM_{rw,diff_{sed}}}{dt}}_{\text{Input and output by sediment diffusion}} + \underbrace{Recharge_{ground}(t)}_{\text{Input from recharge of ground water}} +$$

$$\underbrace{Q_{fish}(t) \cdot Elimination_{fish}}_{\text{Input from elimination of fish}} - \underbrace{Uptake_{fish}(t)}_{\text{Output by fish uptake}} - \underbrace{Irrigation_{rate} \cdot C_{rw}(t)}_{\text{Output by irrigation}} - \underbrace{M_{rw}(t) \cdot \lambda_{degradation,water}}_{\text{Output by degradation in water}}$$

(2.1)

The equation for bottom sediments is given as follows:

$$\frac{dM_{sed}}{dt} = \pm \underbrace{\frac{dM_{rw,sed_{resus,dep}}}{dt}}_{\text{Input and output by sediment deposition and resuspension}} \pm \underbrace{\frac{dM_{rw,diff_{sed}}}{dt}}_{\text{Input and output by sediment diffusion}} -$$

$$\underbrace{M_{sed}(t) \cdot \lambda_{degradation,sediment}}_{\text{Output by degradation in sediment}} \qquad (2.2)$$

Each term in Equation (2.1) and (2.2) represents:

- M_{rw}: chemical mass in raw river water (mg)
- $D_{point\_sources}$: time-dependent flux of pollutant into the freshwater system from point sources (mg d^{-1})
- $C_{rw\_upstream}$: chemical concentration in the upsteram of river (mg m^3)
- C_{rw}: chemical concentration in raw river water (mg m^3)
- $Q(t)$: time-dependent flow rate of river (m^3 s^{-1})
- M_{rw,dry_dep}: chemical mass associated with dry deposition of atmospheric particles in raw river water (mg)
- $M_{rw,wet\_dep,part}$: chemical mass associated with wet deposition of atmospheric particles in raw river water (mg)
- $M_{rw,wet\_dep,gas}$: chemical mass associated with wet deposition of gaseous pollutants (rain dissolution) in raw river water (mg)
- $M_{air,diff\_gas}$: chemical mass associated with diffusion of gaseous pollutant in air (mg)
- $M_{rw,diff\_gas}$: chemical mass associated diffusion of gaseous pollutant in raw river water (mg)
- $M_{rw,wash-off}$: chemical mass associated with soil wash-off in raw river water
- $M_{rw,sed\_resus,dep}$: chemical mass associated with resuspention and deposition between raw water and bottom sediment in raw river water (mg)
- $M_{rw,diff\_sed}$: chemical mass associated with diffusive exchanges of dissolved contaminant in surface water and sediment pore water in raw river water (mg)
- $Recharge_{ground}$: time-dependent recharge of pollutant from ground water (mg d^{-1})
- Q_{fish}: time-dependent fish quantity (mg)
- $Elimination_{fish}$: elimination ratio of fish quantity (d^{-1})
- $Uptake_{fish}$: uptake of pollutant by fish (mg d^{-1})
- $Irrigation_{rate}$: irrigation rate (m^3 d^{-1})
- $\lambda_{degradation,water}$: degradation rate in water (d^{-1})
- M_{sed}: chemical mass in bottom sediments (mg)
- $\lambda_{degradation,sed}$: degradation rate in bottom sediments (d^{-1})

Air compartment: The mass balance equation of a chemical in air compartment is given as follows:

$$\frac{dM_{air}}{dt} = \underbrace{C_{air,out} \cdot V_{wind} \cdot Area_{side}}_{\text{Input from external area}} - \underbrace{C_{air} \cdot V_{wind} \cdot Area_{side}}_{\text{Output from internal area}} + \underbrace{\frac{dM_{total,diffgas}}{dt}}_{\text{Input by diffusion}} - \underbrace{\frac{dM_{air,diffgas}}{dt}}_{\text{Output by diffusion}} -$$

$$\underbrace{\frac{dM_{air,dry_{dep}}}{dt}}_{\text{Output by dry particulate deposition}} - \underbrace{\frac{dM_{air,wet_{dep,part}}}{dt}}_{\text{Output by wet particulate deposition}} - \underbrace{\frac{dM_{air,wet_{dep,gas}}}{dt}}_{\text{Output by gaseous deposition}} \quad (2.3)$$

Each term in Equation (2.3) represents:

- M_{air}: chemical mass in air (mg)
- C_{air_out}: chemical concentration in external atmospheric area (mg m^3)
- V_{wind}: wind velocity (m d^{-1})
- $Area_{side}$: profile area of atomospheric zone (m^2)
- C_{air}: chemical concentraiton in air (mg m^3)
- $M_{total,diff\_gas}$: sum of chemical mass associated diffusion of gaseous pollutant in raw river water, top layer in soil, and fruit/leaf/grain compartments (mg)
- M_{air,dry_dep}: chemical mass associated with dry deposition of atmospheric particles in air (mg)
- $M_{air,wet\_dep,part}$: chemical mass associated with wet deposition of atmospheric particles in air (mg)
- $M_{air,wet\_dep,gas}$: chemical mass associated with wet deposition of gaseous pollutants (rain dissolution) in air (mg)

Soil compartment: It consists of three zones, i.e., surface soil zone, vadose zone, and groundwater zone. The surface soil and vadose zones comprise several layers from 'top layer' to 'bottom layer'. The number of layers can be defined by model users. The mass balance equation of a chemical in surface soil zone is given as follows:

$$\frac{dM_{soil,top}}{dt} =$$

$$\underbrace{\frac{dM_{soil,dry_{dep}}}{dt}}_{\text{Input from dry particulate deposition}} + \underbrace{\frac{dM_{soil,wet_{dep,part}}}{dt}}_{\text{Input from wet particulate deposition}} + \underbrace{\frac{dM_{soil,wet_{dep,gas}}}{dt}}_{\text{Input from gaseous deposition}} +$$

$$\underbrace{\frac{dM_{air,diffgas}}{dt}}_{\text{Input by diffusion}} - \underbrace{\frac{dM_{soil,diffgas}}{dt}}_{\text{Output by diffusion}} - \underbrace{\frac{dM_{soil,wash-off}}{dt}}_{\text{Output by soil wash-off}} + \underbrace{Irrigation_{rate} \cdot C_{rw}(t)}_{\text{Input from irrigation}} -$$

$$\underbrace{M_{soil,top}(t) \cdot D_{transfer}}_{\text{Output by downward diffusion}} + \underbrace{M_{soil,bottom}(t) \cdot D_{transfer}}_{\text{Input by upward diffusion}} - \underbrace{M_{soil,top}(t) \cdot Adv_{transfer}}_{\text{Output by downward advection}} -$$

$$\underbrace{M_{soil,top}(t) \cdot \lambda_{degradation,soil}}_{\text{Output by degradation in soil}} \quad (2.4)$$

The equation for vadose zone is given as follows:

$$\frac{dM_{vadose,top}}{dt} = \underbrace{M_{soil,bottom}(t) \cdot Adv_{transfer}}_{\text{Input by downward advection}} - \underbrace{M_{vadose,top}(t) \cdot D_{transfer}}_{\text{Output by downward diffusion}} + \underbrace{M_{vadose,bottom}(t) \cdot D_{transfer}}_{\text{Input by upward diffusion}} -$$

$$\underbrace{M_{vadose,top}(t) \cdot Adv_{transfer}}_{\text{Output by downward advection}} - \underbrace{M_{vadose,top}(t) \cdot \lambda_{degradation,soil}}_{\text{Output by degradation in soil}} \qquad (2.5)$$

The equation for ground water zone is given as follows:

$$\frac{dM_{groundwater}}{dt} = \underbrace{M_{vadose,bottom}(t) \cdot Adv_{transfer}}_{\text{Input by downward advection}} - \underbrace{Recharge_{ground}(t)}_{\text{Output by recharge of groundwater}} \qquad (2.6)$$

Water balance in root zone is defined as follows:

dW(t)/dt = Precipitation (t) + Irrigation (t) + Groundwater (t) - Evapotranspiration (t) - Percolation (t)   (2.7)

Each term in Equation (2.4) to (2.7) represents:

- $M_{soil,top}$: chemical mass in top layer of surface soil (mg)
- $M_{soil,bottom}$: chemical mass in bottom layer of surface soil (mg)
- $M_{vadose,top}$: chemical mass in top layer of vadose zone (mg)
- $M_{vadose,bottom}$: chemical mass in bottom layer of vadose zone (mg)
- $M_{grondwater}$: chemical mass in groundwater (mg)
- M_{soil,dry_dep}: chemical mass associated with dry deposition of atmospheric particles in top layer of surface soil (mg)
- $M_{soil,wet\_dep,part}$: chemical mass associated with wet deposition of atmospheric particles in top layer of surface soil (mg)
- $M_{soil,wet\_dep,gas}$: chemical mass associated with wet deposition of gaseous pollutants (rain dissolution) in top layer of surface soil (mg)
- $M_{soil,diff\_gas}$: chemical mass associated with diffusion of gaseous pollutant in top layer of surface soil (mg)
- $M_{soil,wash-off}$: chemical mass associated with soil wash-off in top layer of surface soil (mg)
- $D_{transfer}$: transfer coefficient by diffusion (d^{-1})
- $Adv_{transfer}$: transfer coefficient by advection (d^{-1})
- $\lambda_{degradation,soil}$: degradation rate in soil (d^{-1})
- W(t): amount of water storage in root zone (mm d^{-1})
- Precipitation (t): time-dependent precipitation rate (mm d^{-1})
- Irrigation (t): time-dependent irrigation rate (mm d^{-1})
- Groundwater (t): time-dependent groundwater contribution (mm d^{-1})

- Evapotranspiration (t): time-dependent evapotranspiration rate (mm d^{-1})
- Percolation (t): time-dependent percolation rate (mm d^{-1})

Plant compartments (for organic chemicals)

(1) Root compartment

The change of chemical concentration in thick roots can be described as follows: influx with xylem water minus outflux with xylem water, growth and degradation.

$$\frac{C_R}{dt} = \frac{Q}{M_R} K_{WS} C_S - \frac{Q}{M_R K_{RW}} C_R - K C_R \qquad (2.8)$$

where:

- C_R (mg kg fw^{-1}) is the concentration of chemical in root
- Q (L d^{-1}) is the transpiration stream
- M_R (kg fw) is mass of the root
- K_{WS} (kg ww L^{-1}) is the partition coefficient between soil pore water and bulk soil (K_{WS} = C_W / C_S)
- C_S (mg kg ww^{-1}) is the concentration of chemical in bulk soil
- K_{RW} (L kg fw^{-1}) is the partition coefficient between root and water
- k (d^{-1}) is the sum of k_G (d^{-1}) the growth rate and k_{deg} (d^{-1}) the degradation rate

(2) Potato compartment

A potato model that considers diffusion from soil into spherical potatoes, dilution by growth and degradation is applied as follows:

$$\frac{C_P}{dt} = k_1 C_S - k_2 C_p - k C_p \qquad (2.9)$$

where:

- C_P (mg kg fw^{-1}) is the concentration of chemical in potato
- k_1 (d^{-1}) is the uptake rate
- k_2 (d^{-1}) is the depuration rate
- k (d^{-1}) is the sum of k_G (d^{-1}) the growth rate and k_{deg} (d^{-1}) the degradation rate

(3) Leaf compartment

The change of chemical concentration in leaves is influx with transpiration water plus gaseous and particulate deposition from air plus soil attachment minus diffusion to air, growth and degradation.

$$\frac{dC_L}{dt} =$$

$$\frac{Q}{M_L K_{RW}} C_{R,L} + \frac{A_L g_L}{M_L} C_{A,gas} + \frac{(f_{wet}\Lambda_{part}\text{Rain}+f_{dry}v_{dep,dry})S_{field}}{M_L} \text{TSP}_A C_{A,part} + \frac{f_{wet}\text{Irr }S_{field}}{M_L} C_{River} -$$

$$\frac{A_L g_L}{K_{LA} M_L} C_L - kC_L \qquad (2.10)$$

where:

- C_L (mg kg fw^{-1}) is the concentration of chemical in leaves
- $C_{R,L}$ (mg kg fw^{-1}) is the concentration of chemical in influx from root to leaves
- M_L (kg fw) is the mass of leaves
- A_L (m^2) is the area of leaves
- g_L (m d^{-1}) is the conductance of leaves
- $C_{A,gas}$ (mg m^{-3}) is the gaseous concentration of chemical in air
- $C_{A,part}$ (mg g^{-1}) is the concentration of chemical at particles in air
- TSP_A (g m^{-3}) is the total amount of particles in air
- Λ_{part} (m^3 air m^{-3} rain) is the rainfall scavenging ratio for particles
- Rain (m d^{-1}) is rainfall
- $v_{dep,dry}$ (m d^{-1}) is the dry deposition velocity of particles
- S_{field} (m^2) is the surface area of the field
- Irr (mm d^{-1}) is the irrigation rate
- C_{River} (mg L^{-1}) is the concentration of chemical in river water
- ρ_L (kg fw m^{-3}) is the density of leaves
- K_{LA} (m^3 kg fw^{-1}) is the partition coefficient between leaves and air
- k (d^{-1}) is the sum of k_G (d^{-1}) the growth rate and k_{Loss} (d^{-1}) the degradation rate

(4) Grain compartment

The change of chemical concentration in grains is influx with transpiration water plus gaseous deposition from air plus soil attachment minus diffusion to air, growth and degradation.

$$\frac{dC_G}{dt} = \frac{Q_G}{M_G K_{RW}} C_{R,G} + \frac{A_G g_G}{M_G} C_{A,gas} + \frac{f_{wet}\text{Irr }S_{field}}{M_G} C_{River} - \frac{A_G g_G}{K_{GA} M_G} C_G - kC_G \qquad (2.11)$$

where:

- C_G (mg kg fw^{-1}) is the concentration of chemical in grains
- $C_{R,G}$ (mg kg fw^{-1}) is the concentration of chemical in influx from root to grains
- Q_G (L d^{-1}) is the transpiration stream in grain
- M_G (kg fw) is the mass of grains
- A_G (m^2) is the area of grains

- g_G (m d^{-1}) is the conductance of grains. ρ_G (kg fw m^{-3}) is the density of grains
- K_{GA} (m^3 kg fw^{-1}) is the partition coefficient between grains and air.
- k (d^{-1}) is the sum of k_G (d^{-1}) the growth rate and k_{deg} (d^{-1}) the degradation rate

(5) Fruit compartment

The change of chemical concentration in fruits is influx with xylem and phloem plus gaseous and particulate deposition from air minus diffusion to air, growth and degradation.

$$\frac{dC_F}{dt} = \frac{Q_F}{M_F K_{RW}} C_{R,F} + \frac{A_F g_F}{M_F} C_{A,gas} + \frac{(f_{wet}\Lambda_{part}Rain + f_{dry}v_{dep,dry})S_{field}}{M_F} TSP_A C_{A,part} - \frac{A_F g_F}{K_{FA}M_F} C_F - kC_F$$

(2.12)

where:

- C_F (mg kg fw^{-1}) is the concentration of chemical in fruits
- $C_{R,F}$ (mg kg fw^{-1}) is the concentration of chemical in influx from root to fruits
- Q_F (L d^{-1}) is the sum of xylem and phloem flow to fruits
- M_F (kg fw) is the mass of fruits
- A_F (m^2) is the area of fruits
- g_F (m d^{-1}) is the conductance of fruits
- K_{FA} (m^3 kg fw^{-1}) is the partition coefficient between fruit and air
- k (d^{-1}) is the sum of k_G (d^{-1}) the growth rate and k_{deg} (d^{-1}) the degradation rate

Animal compartment:

The change of chemical mass in animals is from input with diet, drinking water and inhalation minus outflux with lipids, urination and exhalation. For milk cattle there is also outflux by lactation. The mass in cattle fat is found by division with the lipid content of the cattle.

$$\frac{dm_C}{dt} = I - km_C \qquad (2.13)$$

where:

- m_C (mg) is the mass of chemical in cattle
- I (mg d^{-1}) is the sum of daily intake of chemical
- k (d^{-1}) is the loss rate constant

The concentration of chemical in milk is found by assuming equilibrium partitioning between milk cattle and milk.

$$C_M = K_{MC}C_c \qquad (2.14)$$

where:

- C_M (mg kg^{-1}) is the concentration of chemical in milk

- K_{MC} (kg kg^{-1}) is the partition coefficient between milk and cattle body
- C_C (mg kg^{-1}) is the concentration in milk cattle found from Equation (2.13) by division with bodyweight of milk cattle and with parameters for milk cattle

The general description of the mass balance equations for human compartment (PBPK model) is described in Chapter 4.

3   PROBABILISTIC PARAMETERIZATION

The general depiction about probabilistic parameterization is presented in the previous section (1.3). The focus of this chapter is set at the detailed presentation of two advanced statistical approaches to derive probabilistic density function (PDF) for the two parameters used in the freshwater compartment of the 2-FUN multimedia model: the settling velocity of particles that is a driving factor influencing the transfer of particles at the water-sediment interface in fresh-water system, and the fish bioconcentration factor (BCF) for metal that represents the accumulation of a given chemical in organisms arising by water uptake. The third section summarizes the other approaches used to estimate probabilistically the parameters associated with chemical transfer between water and sediment phases in the freshwater compartment. The detailed description for it is available in a deliverable of 2-FUN project and Ciffroy et al (2010b).

3.1   *Probabilistic estimation of the settling velocity of particles using a Bayesian approach*

3.1.1   Introduction

The deposition and resuspension of particles occurring at the water-sediment interface in fresh-water systems are the main transfer processes concerning the exchange of contaminants between the water phase and the bottom sediment. These transfer processes are commonly considered as key processes for the environmental risk assessment of chemicals because these processes strongly affect the residence time of contaminants associated with particles in fresh-water.

In most existing multi-media models such as SimpleBox (Brandes et al. 1996), QWASI (Mackay 1983; Di Guardo et al. 2006; Warren et al. 2007) and Trim.Fate (USEPA 2002), the processes of deposition and resuspension are each represented by a steady-state rate and expressed as the downwards and upwards velocities of particles, occurring simultaneously. Thus,

the mass balance of particles at the water-sediment interface in the steady-state scheme is defined by the net sedimentation rate which is calculated as the deposition rate minus the resuspension rate. The main drawback of this steady-state representation is that it assumes a permanent net deposition rate when the deposition rate is greater than the resuspension rate. This assumption may be not realistic for many freshwater systems where there are temporal cycles of deposition/resuspention caused by the seasonal variations in the flow rate.

Mechanistic time-dependent deposition/resuspension equations have been proposed for use in hydraulic models (e.g. Krone 1962; Partheniades 1965; Blom and Aalderink 1998; Sanford and Maa 2001; Winterwerp 2006) and can be assumed to be a more realistic representation of the physical processes. Critical analysis of such equations is beyond the scope of our study, however, Ciffroy et al (2010a) evaluated these equations and selected the simultaneous deposition/resuspension model recommended by Winterwerp (2006). We focus *a priori* on these mechanistic time-dependent equations which are based on the parameters having physical meaning such as the settling velocity of particles, the critical shear stress of deposition/resuspention, and the erosion rate, while remaining mathematically simple. Therefore these equations can be incorporated into the scheme of multi-media models.

In those equations accounting for the deposition process described in Krone (1962) and Blom and Aalderinki (1998), the settling velocity of particles is defined as one of the key parameters in transferring contaminants from the water phase to the bottom sediment (Droppo et al. 2000). The parameter values are related to the natural conditions such as particle size and density, and therefore tend to be highly variable and site specific. Especially in the context of multi-media modelling, using a representative (single) value of the parameter for a fresh-water system such as river, lake or pond, regardless of site-specific natural conditions, is likely to ignore the potential parametric variation seen among different sites.

A probabilistic representation of the settling velocity of particles should instead be preferred to capture the parametric uncertainty with which values of the settling velocity occur in the natural environment. A PDF for the settling velocity of particles can be estimated from two types of knowledge as follows;
- A theoretical formula, taking into account some physical characteristics of particles (i.e. size and density).
- Measurements taken from the literature.

Neither of these approaches can be considered as 'the best' approach because they both present advantages but also drawbacks (discussed in the next section).The aim of our study was to

derive an improved generic (global) PDF for the settling velocity of particles by a statistical method which integrates two types of information. Furthermore, our study also aimed to demonstrate how the statistical method can update the estimated generic PDF, when new site-specific measurement data of the settling velocity are incorporated into the generic PDF.

3.1.2 Materials and methods

Estimation of the settling velocity of particles by a simple theory

When the water flow around the particles is laminar (not turbulent) and particles are not subject to aggregation under low suspended concentrations, the settling velocity of particles W_c can be described by Stokes' law. The equation describing the settling velocity of small spheres (diameter <0.1 mm) with uniform density in the particle Reynolds number regime (Mantovanelli et al, 2006) is as follows:

When $\mathrm{Re} = \dfrac{W_c \cdot d \cdot \rho_w}{\mu} < 1$ then,

$$W_c = 86400 \cdot \frac{g \cdot d^2 \cdot \rho_e}{18\mu} \quad (3.1)$$

All the variables in Equation (3.1) are shown with their values in Table 3.1.

Table 3.1: Variables and parameters used in Equation (3.1)

Parameter	Description of parameter	Given value and unit
W_c	Settling velocity of particles	m d^{-1}
ρ_w	Water density	998 kg m^{-3}(at 20°C)
d	Diameter of particles	m
ρ_e	Excess density (or effective density) of particles, equivalent to $\rho_e = \rho_s - \rho_w$ where ρ_s is particle density. The particle density was set to the average value (2300 kg m^{-3}) typical of clay minerals as seen in Khelifa and Hill (2006).	2300 kg m^{-3}
μ	Dynamic molecular viscosity of water	10^{-3} kg m^{-1} s^{-1}(at 20°C)
Re	Particle Reynolds number	-
υ	Kinematic viscosity of water	10^{-6} m^2 s^{-1}(at 20°C)
g	Gravity acceleration	9.8 m s^{-2}

In the natural aquatic environment, fine-grained particles tend to aggregate into larger porous aggregates, known as flocs, due to particle collisions mainly caused by Brownian and turbulent motion, and surface electro-chemical forces (Winterwerp, 1998; Liu et al. 2002). As a result of this aggregation process, cohesive sediments settle as flocs rather than as individual particles (Liu et al. 2002). Various experiments indicate that large suspended particles that emerge as aggregates of much smaller particles play a key role in determining the fate of fluvial contaminants such as heavy metals and pesticides (Lick and Rapaka, 1996; Leppard et al. 1998).Therefore, analyzing the settling of flocs is essential to understand the transfer of contaminants from the water phase to the bottom sediment.

Two models (Khelifa and Hill, 2006; Winterwerp, 1998) have been proposed to estimate the settling velocity of flocs. These models are based on the concept of floc fractal dimension F, and it was demonstrated that they matched the observed data adequately. However, these models require the derivation of the key parameter F from the relationship between the measurements of excess floc density and floc size. The processes of parameterization of the fractal dimension F and estimation of the settling velocity of flocs are discussed in Fettweis (2008). Furthermore, estimation of the particle Reynolds number Re in these models seems to require a series of measurements of the settling velocity.

Calculation of the settling velocity of flocs by Stokes' law (Equation (3.1)) entails a certain inaccuracy, as Stokes' law was derived without consideration of turbulence and the fractal geometry of flocs. Nevertheless, we selected Stokes' law as a theoretical model to estimate the settling velocity because: (i) in two studies (Liu et al. 2002; Fox et al. 2004), Stokes' law was used to estimate the parameter under natural conditions, and (ii) in the scheme of our study, the parameter values roughly estimated by Stokes' law are defined as *prior* information and the inaccuracy contained in the prior information is assumed to be compensated for by combining this information with the other information from the measurements. In addition, an advantage of using this model is that it requires consideration of only two variables viz the diameter and density of flocs.

Based on Stokes law, a Monte Carlo approach was applied to build a theoretical prior PDF of settling velocities, considering a range and distribution of particle sizes and a typical density of particles. In the Monte Carlo calculation, the particle diameter d was selected as the input parameter with a variable value, whereas other parameters were given fixed values, because the particle size is a variable to which the settling velocity could have great sensitivity. The range of particle sizes (flocs) in natural aquatic conditions ranges over three to four orders of magnitude

(Lartiges et al. 2001; Mikkelsen et al. 2005) and can be really site-specific, and thus defining a full range of particle sizes remains an open question. In constructing a theoretical prior PDF for this study, we focused on the range of the smaller particle sizes because it can be assumed that the presence of a lower range of particle sizes is not site-specific but is commonly seen in any natural conditions, whereas the higher range is highly variable depending on the natural conditions at each location. An upper value of the range of d was set at 63 μm based on a finding by Sverdrup et al. (1942) that Stokes' law can be considered valid for spheres of diameter <62.5 μm in quiescent waters. The bottom of the range was set at 1 μm because, unlike submicron colloidal particles, suspended particles >1 μm should be recognized as a significant vector for the transport of anthropogenic contaminants in surface waters (Lartiges et al. 2001). According to some scientific reports, the particle-size distribution of a finely divided system is often found to be log-normal (Kiss et al. 1999; Wagner and Ding 1994; Kim et al. 1995). Furthermore, in sedimentary research, the log-normal model for grain-size distribution was developed by Krumbein (1938), and empirical evidence and a possible explanation have been offered for this model (Purkait 2002; Ghosh 1988; Clarke and Ghosh 1995). This log-normal model for the particle-size distribution may not fully match the realistic distribution considering the presence of flocs in the natural environment. Nevertheless, we assigned a log-normal distribution to the particle size in order to simplify the parameterization of the variable. A log-normal distribution for the variable was built by arbitrarily defining the range between 1 and 63 μm as the 99.8 % interval of the distribution, i.e. setting these values at the 0.1 and 99.9 percentile, respectively.

Collection of experimental and in-situ measurements of the settling velocity of particles

Measurements of the settling velocity of particles were obtained from the literature review. For each such experiment or *in situ* measurement, we present a representative measured value of the parameter, the range of the observed particle diameters and a short description of the experimental method employed (Table 3.2). These measurements were taken from differing aquatic systems such as an experimental flume, river, estuary, and gulf. In our study, we focused on the settling velocity observed in fresh-water systems, especially rivers. However, due to the scarcity of measurements for riverine systems and the aim to derive a 'generic' PDF, we decided to aggregate all the data from differing aquatic systems in order to give a certain parametric uncertainty which must be taken into consideration in the natural environment.

The experimental and measurement approaches seen in the literature review are mainly categorized into three types: (i) laboratory flume experiments (Beaulieu et al. 2005; El Ganaoui et al. 2004; Graham and Manning 2007), using channels where sediments can be deposited and where initiation of resuspension is observed under increasing stream velocity; (ii) indirect *in situ*

measurements, where parameters are fitted from suspended particulate matter (SPM) monitoring in surface water (Blom and Aalderink 1998; Luck 2001; Voulgaris and Meyers 2004), and (iii) direct *in situ* measurements with sediment traps or other sensors (Graham and Manning 2007; Kozerski 2002). In total, 13 values were obtained for the parameter (Table 3.2). When more than one value of the settling velocity was found in each paper or each different experimental and measurement method, their mean values were taken in order to give an equivalent weight to each source of information.

Table 3.2: Database of experimental and in-situ measurement for the settling velocity of particles

Reference	Settling velocity of particles – W_c (m d^{-1})	Particle size (μ g)	Experimental method
Beaulieu et al. (2005)	a. 11.23 (rough estimate) b. 25.92 (rough estimate at flow speed 3 cm s^{-1})	No data in this study, but a range 1 – 100 is cited from Archambault et al. (2003) which made similar flume experiments	Two types of flume experiments considering clay/algal flocculation a. Straight-channel 17 m flume experiment (without flow speed) b. Racetrack flume experiment (with flow speed)
Blom and Aalderink (1998)	a. Not estimated because of resuspension conditions b. Mean value 19 [range: 13-26]		a. Flume experiment with natural sediments from Lake Ketel (The Netherlands) – A stepwise increasing bottom shear stress was applied. b. Six hourly *in situ* monitoring of SPM concentrations over 1 week
Ciffroy et al. (2000)	34.56, 25.92, 25.92, and 8.64 (These values are used only for calculation of the site-specific PDF)		Measurements of SPM vertical gradients and application of the Rouse profile theory (The Seine river, France)
Curran et al. (2007)	15.5		Estimation of particle settling velocity by the measurement of particle mass distributions using an *in situ* column tripod equipped with floc camera, video camera, and laser scatterometer (Gulf of Lions, France)
De Jesus Mendes et al. (2007)	Mean value 13.25 [range: 12.96-13.82]	28 - 213	Measurement of the settling velocity of sampled organo-mineral aggregates in a container maintained under *in situ* temperatures and using a settling cylinder with

			a video camera (sampling place: Setùbal Canyon, Portugese west continental margin)
El Ganaoui et al. (2004)	a. Mean value: 2.43 [range: 0.99-3.89] b. Mean value: 5.76 [range: 4.32-7.78]		Flume experiment with sediments from three stations of the Rhône river and the Gulf of Lion (France) – Two-class model with a. easily erodable particles (fluff layer) b. classical cohesive particles
Fettweis (2008)	Mean value 6.45[range: 0.26-17.28[1]]	44 - 160	Estimation of the excess density and the settling velocity of mud flocs by measurement using an optical sensor, SPM siltration and laser scatterometer (Belgian near-shore zone)
Graham and Manning (2007)	Mean value 47.52 [range: 0.34 -290.3]	20 - 1265	Flume experiments with macrophytes - Measurements of full spectral floc size and particle settling velocity variability using floc camera
Hill et al. (2000)	Mean values 7.43 [range: 5.18-8.64[2]]	370 ± 100 (mean and standard deviation)	Estimation of the effective settling velocity by along-shelf measurements of water salinity and sediment concentration within t Niskin bottles and by point measurements of plume velocity (sampling place: Eel river mouth and along-shore points, California, USA)
Kozerski (2002)	4.3		Use of plate sediment traps deployed in the Spree river and evaluation of trapping rates (in gDW m^{-2} d^{-1})
Luck (2001)	1		Long-term monitoring of SPM and calibration of a hydro-sedimentological model (The Loire river, France)
Voulgaris and Meyers (2004)	Mean value 23.62[range: 8.64-31.97]	25 - 75	*In situ* measurements of SPM concentration and particle-settling velocity variability using laser scatterometer, optical sensors, and Doppler velocimeter (Bly Creek, South Carolina, USA)

[1]: Five mean tidal values with relatively low standard deviation (± 0.015 – 0.18) were chosen and averaged

[2]: The five effective settling velocities were chosen and averaged, which are significantly different from zero at the 95% confidence interval.

In many cases, the parametric PDFs have been derived only from measurements, even if they are scarce and often contain measurement errors, for the reason that measurements can give 'real' values in contrast to the 'estimated' values obtained from theoretical considerations. However, when the measurement information is poor, the variation that measurements contain may not cover the potential variation seen under natural conditions. The settling velocity measured in the considered literature is associated with the following ranges of particle size (in diameter): (i) 28 to 213 μm (De Jesus Mendes et al. 2007), (ii) 44 to 160 μm (Fettweis 2008), (iii) 20 to 1265 μm (Graham and Manning 2007), (iv) 370 $\pm$ 10 μm (mean and standard deviation) (Hill et al. 2000), and (v) 20 to 75 μm (Voulgaris and Meyers 2004). These ranges of particle size are higher than the range we assigned, 1 to 63 μm, for estimating a theoretical prior PDF. Considering that the settling velocity generally increases with increasing particle size, the settling velocity measured under conditions with the particle sizes described above probably takes the higher range of values than that of the theoretical values. It is thus important to combine both the theoretical information and the measurements in order to allow coverage of non-measured situations, i.e. to take into account the potential variation of parameter values as much as possible.

A Bayesian approach to combine the theoretical and measurement information

A Bayesian approach was selected to combine the two sources of information because it is well adapted to build a PDF when theoretical prior knowledge is available and when only a limited set of measurements can be collected for a parameter. This approach calculates the posterior distribution by assuming a log-normal measurement model with unknown parameters μ and σ^2 (representing the mean and variance of logarithmic data, respectively), knowing a prior distribution (here derived from Stokes' law) and given the measurements (Data) of the parameter. According to Bayes' theorem, the posterior PDF is given by:

$$p(\mu, \sigma^2 | \text{Data}) = \frac{p(\text{Data}|\mu, \sigma^2) \cdot p(\mu, \sigma^2)}{p(\text{Data})} \qquad (3.2)$$

On the right side of Equation (3.2), the parameter $p(\text{Data} | \mu, \sigma^2)$, also known as the data likelihood, is the probability of observing the data vector conditional on the parameter vector (μ, σ^2). The function $p(\mu, \sigma^2)$ does not involve the data vector; it represents the uncertainty of the parameter vector (μ, σ^2) before data are observed and is called the prior distribution of (μ, σ^2). The denominator is a function of 'Data' which is observed and is completely known. In practical calculations, all constant factors can be ignored and thus Bayes' theorem can be re-written using proportionality as:

$$p(\mu, \sigma^2 | \text{Data}) \propto p(\text{Data}|\mu, \sigma^2) \cdot p(\mu, \sigma^2) \qquad (3.3)$$

This theorem updates a given prior distribution with information contained in measurements and gives a posterior distribution which corresponds to the required PDFs. The successive steps of the Bayesian approach are illustrated in the next sections.

Derivation of a generic PDF of the settling velocity of particles by a Bayesian approach

The following procedure was applied to the settling velocity of particles (hereafter W_c) in order to build its posterior distribution. A theoretical prior PDF for W_c estimated by Stokes' law can be described as $\ln(W_c) \sim N(\mu_{FC}, \sigma_{FC}^2)$ where N denotes normal distribution, and μ_{FC} and σ_{FC}^2 represent respectively the logarithmic mean and variance of the theoretical prior PDF. In our approach, these parameters are regarded as the mean and variance of a 'fictional' sample from a normal distribution, and used to provide *a priori* information for both the logarithmic mean μ_{Wc} and variance σ_{Wc}^2 of W_c. In general, the prior distribution obtained from such definite information is called informative prior. Using the information of μ_{FC} and σ_{FC}^2, the prior distribution for μ_{Wc} is expressed as $N(\mu_{FC}, \sigma_{Wc}^2/n)$ where N indicates a normal distribution with mean μ_{FC} and variance σ_{Wc}^2/n. On the other hand, the prior distribution of σ_{Wc}^2 is expressed as $(n-1) \cdot \sigma_{FC}^2/\sigma_{Wc}^2 \sim \chi^2(n-1)$ where $\chi^2(n-1)$ denotes the Chi square distribution with n-1 degrees of freedom. In this approach, n is defined as the fictional number of observations and given by the actual number of measurement data (i.e. n = 13) in order to give an equivalent weight ('belief') to both the theoretical and measurement information of W_c for calculation of its posterior distribution. The reason to give them an equivalent weight is that they can potentially cover different ranges of parameter values, and there is no rationale to evaluate differentially the qualities of the two sets of information.

In other words, the *a priori* knowledge of the distribution from which each value of μ_{Wc} is assumed to be drawn is represented by a normal distribution centered around the sample mean and scaled by the number of measurements, and the distribution from which each value of σ_{Wc}^2 is assumed to be drawn is a Scaled Inverse Chi Square distribution centered around the sample variance and with n-1 degrees of freedom.

Bayes' theorem (Equation (3.3)) incorporating the collected dataset and the prior distributions $p(\mu_{Wc})$ and $p(\sigma_{Wc}^2)$ defines the joint (i.e. two-dimensional) posterior distribution $p(\mu_{Wc}, \sigma_{Wc}^2 | \ln(y))$ of μ_{Wc} and σ_{Wc}^2, where $\ln(y)$ denotes the natural logarithm of the dataset (measurements). Simulation from this distribution was performed using the Gibbs Sampler algorithm (Gelfand et al. 1990) programmed in MATLAB using BABAR® software (www.facilia.se). Briefly, the algorithm draws μ_{Wc} and σ_{Wc}^2 from their two-dimensional posterior distribution by drawing iteratively from the one-dimensional conditional posterior

distribution $p(\mu_{Wc}|\sigma^2_{Wc}, \ln(y))$ and $p(\sigma^2_{Wc}|\mu_{Wc}, \ln(y))$ respectively, until the sampling reaches convergence.

Given draws of μ^k_{Wc} and $\sigma^{2,k}_{Wc}$, the posterior predicted distribution of W_c was then simulated by drawing each value of W_c^k k=1,...,n (n = the number of samples: 50,000 in the present study) from the log-normal distribution $LN(\mu^k_{Wc}, \sigma^{2,k}_{Wc})$ where each μ^k_{Wc} and $\sigma^{2,k}_{Wc}$ is taken as the k_{th} value obtained from the joint posterior distribution $p(\mu_{Wc}, \sigma^2_{Wc}|\ln(y))$.

Derivation of a site-specific PDF of the settling velocity of particles by a Bayesian approach

Since the generic PDF of W_c was estimated updating the theoretical prior PDF with the aggregation of measurement data which represents a 'generic' parametric variation seen among differing natural conditions, the uncertainty contained in the generic PDF tends to be high. Due to the potentially high uncertainty, the generic information of the parameter can be suitable for modelling studies dealing with a generic (regional or continental) scenario but not for modelling studies dealing with a site-specific (local) scenario. Therefore, in our study, a site-specific PDF of W_c was derived utilizing knowledge of the generic PDF and site-specific measurements, aiming at observing the effect on the uncertainty contained in the generic PDF. The Bayesian approach updated the generic PDF which is regarded as prior information with site-specific measurements of the parameter, and estimated a site-specific PDF as a posterior distribution.

We selected a different way for prior setting in the Bayesian approach to emphasise our 'belief' in the information from site-specific measurements containing substantive information associated with a specific location as opposed to the *a priori* knowledge given by the generic PDF potentially containing high variation associated with multiple locations. In this method, the logarithmic mean μ_{ge} and logarithmic variance σ^2_{ge} of the generic PDF give *a priori* information as an informative prior only for the logarithmic mean μ_{Wc} of W_c. Furthermore, the *a priori* knowledge of the distribution from which each value of μ_{Wc} is assumed to be drawn can be represented by a normal distribution with hyper-parameters μ_0 and σ^2_0 (i.e., $\mu_{Wc} \sim N(\mu_0, \sigma^2_0)$). Under the interpretation that the generic PDF is representing *a priori* knowledge of probable values of μ_{Wc}, the hyper-parameters were chosen as $\mu_0 = \mu_{ge}$ and $\sigma^2_0 = \sigma^2_{ge}$. For the logarithmic variance σ^2_{Wc} of W_c, on the other hand, no such definite *a priori* knowledge can be extracted from the theoretical prior PDF. One non-informative prior is obtained by assigning an unbounded uniform prior distribution for σ_{Wc} on logarithmic scale. That is $\ln(\sigma_{Wc}) \sim Uniform(-infinity, infinity)$ which can be shown to be equal to using $p(\sigma^2) \propto 1/\sigma^2$ in Equation (3). This distribution is sometimes used as a non-informative prior for the variance parameter of normal models when the data likelihood provides substantive information

(Detailed description is in Gelman et al. 2004). Subsequent procedures to estimate the posterior predicted distribution of Wc were carried out in the same manner.

Local measurements of W_c were taken from Ciffroy et al (2000). In the study, the local measurements were monitored at four sampling points located upstream and downstream in the Seine River in France. The four best estimates of the parameter were 34.6, 25.9, 25.9 and 8.6 m d^{-1} for each sampling point.

3.1.3 Results and Discussion

A theoretical prior distribution for the settling velocity of particles

The generation of a random sample from the log-normal distribution of the particle diameter d was conducted by a Latin Hypercube Sampling (LHS) scheme because it gives better coverage of the parameter space than random Monte Carlo sampling (Helton 1993). The forecast values of W_c were generated by 50,000 runs. Their range meets the condition of a laminar-flow regime which is described by the Reynolds number Re < 1. A log-normal distribution was fitted to forecast values of W_c and chosen as a theoretical prior PDF of W_c where $p(\ln(W_c)) \sim N(\mu_{Fc}: 1.39, \sigma_{FC}^2: 1.72)$. On the normal scale, these μ_{FC} and σ_{FC}^2 can be converted into geometric mean (GM) and geometric standard deviation (GSD) as $\exp(\mu_{FC}) = GM, \exp\left(\sqrt{\sigma_{FC}^2}\right) = GSD$. The given GM and GSD representing the theoretical prior PDF on a normal scale are 4 and 3.7, respectively. The corresponding values for the 13 measurements are 9.3 and 2.9. They were calculated as $GM = \exp(Mean(\ln(y)))$, $GSD = \exp(SD(\ln(y)))$, where SD and y denote standard deviation and measurements, respectively. From their GM values, the value of measurements is more than twice as high as that of the theoretical prior PDF.

A generic PDF of the settling velocity of particles (1st simulation)

The joint (two-dimensional) posterior distribution of μ_{Wc} and σ_{Wc}^2 for settling velocity of particles (W_c) was defined by Bayes' equation which combines the obtained prior distributions for μ_{Wc} and σ_{Wc}^2 and experimental data for W_c. Values from the posterior distribution for 50,000 samples were generated by the Gibbs sampling method. Mean values of simulated posterior estimates of μ_{Wc} and σ_{Wc}^2 are 1.81 and 1.71, respectively. The prediction of posterior estimates of W_c was performed based on both μ_{Wc} and σ_{Wc}^2 posterior distributions as described in the previous section.

The GM and GSD representing the distribution of posterior predicted samples for W_c were calculated as $GM = \exp(Mean(\ln(W_{c\_}pred)))$, $GSD = \exp(SD(\ln(W_{c\_}pred)))$, where

W_c_pred denotes predicted samples of W_c. Note that the distribution of sampled W_c estimates approximately follows log-normality. Table 3.3 presents all the geometric means (GMs) and geometric standard deviations (GSDs) of the distributions for posterior predicted samples, the theoretical prior and the measurements.

A site-specific PDF of the settling velocity of particles (2nd simulation)

The generic PDF estimated as a posterior distribution was used with the logarithmic form, μ_{ge} and σ_{ge}^2, in the simulation procedure, and then updated with site-specific measurement data in order to calculate a site-specific PDF for W_c. Mean values of simulated posterior estimates of μ_{Wc} and σ_{Wc}^2 are 2.94 and 1.21, respectively. The distributions of posterior predicted samples, prior given by the generic PDF and site-specific measurements for W_c are presented as their GM and GSD values (Table 3.3).

Table 3.3: Statistical descriptors of prior PDFs, data sets and posterior PDFs for W_c

	Geometric mean (GM) (m d^{-1})	Geometric standard deviation (GSD) (m d^{-1})	Coefficient of variation (CV)
1st simulation for estimating a generic PDF			
Theoretical prior PDF (log-normal)	4	3.7	0.94
Measurements	9.3	2.9	0.48
Posterior distribution (log-normal) : (Generic PDF)	6.1	3.8	0.74
2nd simulation for estimating a site-specific PDF			
Generic PDF as prior distribution	6.1	3.8	0.74
Site-specific measurements in Seine River (France)	21.2	1.8	0.19
Posterior distribution (log-normal) : (Site-specific PDF)	19	3	0.37

Comparison reveals that the GMs of both generic and site-specific PDFs of W_c (Table 3.3) take values somewhere between the values of their corresponding prior distributions and the measurements. In the simulation for the generic PDF (1st simulation), the absolute difference between GMs of the prior distribution and of the posterior distribution is 2.1 whereas the difference between GMs of the measurements and of the posterior distribution is 3.2. For simulation of the site-specific PDF (2nd simulation), on the other hand, the corresponding absolute differences are 12.9 and 2.2. From the 2nd simulation, the GM of the posterior distribution was simulated much more closely to that of the measurements than the prior distribution, whereas GM of the posterior distribution in the 1st simulation was estimated

roughly in the middle of GMs for the global measurements and the theoretical prior distribution. This is because in the 1st simulation, the prior knowledge obtained from Bayes' theory contributed to deriving the posterior distribution almost as significantly as the measurement information. In the 2nd simulation, however, the prior knowledge coming from the generic PDF is not weighted with the number of measurements and thus the contribution for calculating the posterior distribution was relatively weighted toward the site-specific measurement information. All the distributions were drawn based on their GM and GSD values (Figs. 3.1a and 3.1b) to visualize how the distributions are located.
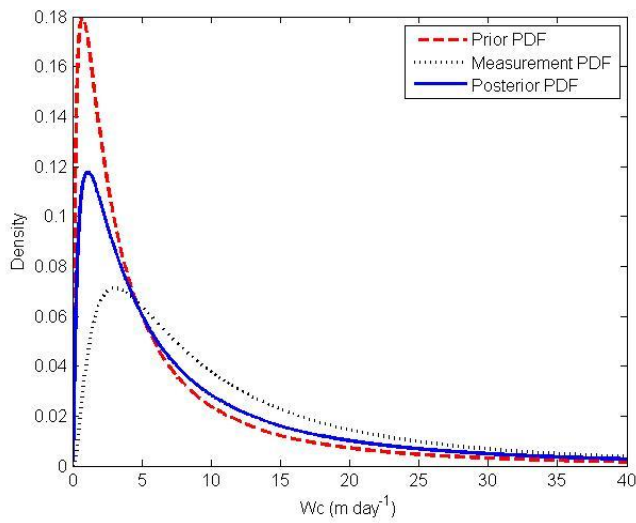

Figure 3.1a: Distributions of posterior predicted samples, theoretical samples and measurements (for W_c 1st simulation for estimating a generic PDF)
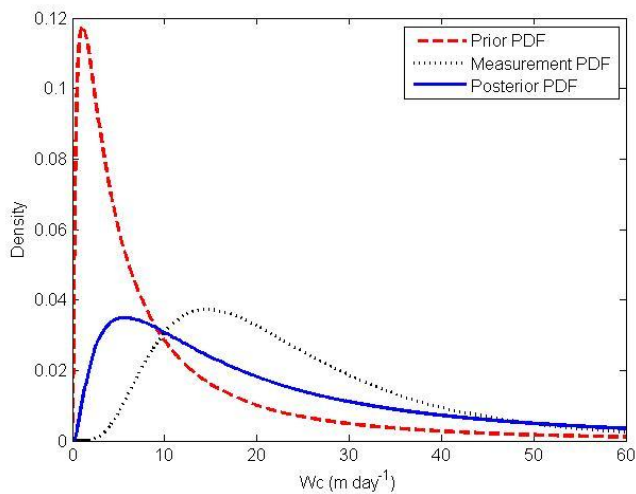

Figure 3.1b: Distributions of posterior predicted samples, theoretical samples and measurements (for W_c 2nd simulation for estimating a site-specific PDF)

The coefficient of variation (CV) was used to compare the degree of variation that each distribution contains. This statistic is generally defined as the ratio of the standard deviation to the mean; thus in this study, the CV was calculated in the form of the ratio of the logarithmic standard deviation ($\sigma = \ln(GSD)$) to the logarithmic mean ($\mu = \ln(GM)$) for each distribution (Table 3.3). Comparison of CVs shows the variations of both generic and site-specific PDFs of W_c become lower than those of the corresponding prior distributions because of the incorporation of measurement information which contains less variation than prior distributions. Especially, in the 2nd simulation, the variation was drastically reduced from 0.74 in the prior distribution to 0.37 in the posterior distribution. This is because the incorporated site-specific measurements have much lower variation (0.19), i.e., give much more informative knowledge on W_c, compared with the prior distribution, and also because the Bayes' approach in the 2nd simulation gave more importance to the measurement information than to the prior information; thus the posterior distribution can get close to the measurement information in terms not only of its GM but also its variation. The posterior distribution comes close to the measurement distribution (Figure 3.1b), while being away from the prior distribution. In the 1st simulation, variation for the posterior distribution was estimated almost in the middle of two variations for prior distribution and measurement data, as seen in the case of their GMs (Figure 3.1a).

3.1.4 Conclusion

The proposed mechanistic equations consist of parameters having physical meaning and represent time-dependent (dynamic) sediment behavior. Thus these equations characterize the physical and natural phenomena at the water-sediment interface better than the current multimedia-modelling scheme assuming steady-state sediment behavior, and can be used in a new multimedia-modelling scheme. When the parameters are set probabilistically in the form of PDF, the mechanistic equations can provide a range of output taking into account parametric uncertainties. In order to grasp realistically the sediment behavior by modelling, it is important to consider both the temporal variation of sediment behavior by mechanistic expressions and the parametric uncertainties by PDFs.

Our study focused on the derivation of the parametric uncertainty of the settling velocity of particles in the form of PDF. A Bayesian method was found to be useful to derive updated probabilistic density functions (PDFs) of the investigated parameter W_c, by integrating both their theoretical and measured information, when neither information is assumed to cover all the potential variation that the parameter may contain. Therefore, the simulated posterior distributions of W_c can be assumed to be more comprehensive PDFs than either prior or measurement information.

The generic or site-specific PDF of W_c can be used in different ways, depending on the modelling scheme to be chosen. When a modelling study is conducted on a large target area with a continental or regional scenario, the generic PDF can be used because it contains a relatively high and 'comprehensive' parametric uncertainty derived from the general theoretical assumption and also from global measurement information. On the other hand, when a small area with a local scenario is considered, i.e. the Seine River in our case, the site-specific PDF can be more practical than the generic PDF because it is more specialized for local measurement information than for generic prior information.

In the present study, a Bayesian approach was used to derive posterior generic and site-specific PDFs for the settling velocity of particles. For further research, it will be necessary to consider all the parametric uncertainty contained in the critical shear stress of deposition/resuspension and in the erosion rate, which are the parameters associated with the dynamic deposition/resuspension equations. Through the literature review, some measurements of those parameters have been located. Thus, potentially, the Bayesian approach is applicable to estimate posterior PDFs of those parameters if theoretical models for those parameters are established.

3.2 *Regression approaches to derive generic and fish group-specific probability density functions of bioconcentration factors for metals*

3.2.1 Introduction

The bioconcentration factor (BCF), especially that for fish, is a parameter commonly considered in environmental multi-media models dedicated to environmental and health risk assessments of chemicals, such as Caltox (Mckone 1993a), QWASI (Mackay et al. 1983; Di Guardo et al. 2006; Warren et al. 2007), and Trim.Fate (U.S.EPA 2002). It represents the ratio, at equilibrium, of a chemical concentration in an organism to the concentration in water and is used in single-compartment models. Bioconcentration factor describes the accumulation of a given chemical in organisms arising by uptake from water only and is generally determined under laboratory conditions, whereas the Bioaccumulation Factor (BAF) represents accumulation from both water and diet and is derived from in situ observations.

The BCF concept was originally developed for hydrophobic organic substances and several quantitative structure-activity relationship (QSAR) techniques were proposed to predict BCF from chemical descriptors of hydrophobicity such as octanol-water partition coefficients (Barber 2003; Schüürmann et al. 2007; Zhao et al. 2008). Simple passive diffusion across the lipid biomembranes is believed to be the key process for the accumulation of neutral hydrophobic

substances in biota, which ensures that BCF is independent of exposure concentrations (McGeer et al. 2003). In the case of metals, however, the assumption of BCF being independent of exposure concentrations was found to be erroneous. As a result of complex physiological processes such as sequestration, detoxification, storage and branchial elimination, biota are actively able to regulate metal bioconcentration via dynamic reaction systems that respond to environmental loading and maintain homeostasis (Hamilton and Mehrle 1986; Chapman et al. 1996). Based on the factors influencing metal uptake and accumulation, it can be assumed that BCF values for metals are not independent of exposure. In addition, DeForest et al. (2007) hypothesized the trend by which metal uptake increases at lower exposure concentrations, according to the basis that organisms actively uptake essential metals at low concentrations to satisfy metabolic requirements. Non-essential metals would also be regulated because the mechanisms for regulating essential metals are not metal-specific (Phillips and Rainbow 1989).

The variation of BCF values for metals (Cd, Cu, Pb, Hg, Zn, As, etc) reported in the literature (McGeer et al. 2003; DeForest et al. 2007) was found to range over several orders of magnitude. The linear regression analysis in those studies observed inverse relationships between BCF values and aquatic-exposure concentrations for various aquatic species and metals. The United States Environmental Protection Agency (U.S.EPA) underlined that the current science does not support the use of a single, generic BCF or BAF value as an indicator of metal hazard for regulatory applications and made some recommendations to risk assessors for alternative approaches for using the BCF concept (U.S.EPA 2007). In particular, the U.S.EPA suggested that risk assessors use regression relationships between tissue and exposure concentrations instead of single values such as mean or median values. Besides, we can presume that the trend in BCF values significantly depends on fish species. For example, McGeer et al. (2003) observed that the slopes of the BCF-exposure regression line vary among different fish species for each metal. Therefore, it can be a practical approach for the risk assessment relevant for the specific investigation site and fish group to use an independent BCF for each specific fish group, i.e., fish family or fish species.

When referring to risk assessment, one must distinguish between screening level assessments where generic and/or worst-case scenarios can be considered and detailed level assessments where more realistic scenarios should be taken into account. Concerning detailed level assessments, the importance of probabilistic techniques, such as a Monte Carlo analysis, to the realistic characterization of environmental risk has been well recognized, especially in the studies using environmental simulation models (Solomon et al. 2000; Dubus et al. 2003; Gutiérrez et al. 2009). In the process of stochastic approaches by the simulation models,

assigning appropriate distributions to input parameters, e.g., the BCF for metals, is a key step in probabilistic analyses as it influences the final results of simulation, i.e. the potential levels of environmental risk for target substances. The distributions assigned to input variables can be described by probability density function (PDF). When the uncertainty contained in the BCF-exposure regression relationship is statistically analyzed, the PDF of BCF can be derived at given exposure metal concentrations. In modelling studies dedicated to the environmental and health risk assessment, the PDF will be used as an input to the probability analysis (such as a Monte Carlo analysis) conducted in the model simulation. Then the result of the model simulation will be expressed by an integrated distribution of potential risk levels for the metal and allow risk assessors to quantify the uncertainty associated with model prediction and also conduct the sensitivity analysis to know how much the parameter (BCF) will influence the result such as the potential risk level of the metal.

The objectives of the present study were, then, to examine the regression model that correlates collected data of fish BCF and the corresponding exposure concentrations for each metal; and to derive the PDF of fish BCF for each metal at a given exposure level in order to test how much the data variation can be reduced by taking into account the uncertainty contained in the regression model. To build the regression model and derive the PDF of fish BCF, we applied two different statistical approaches: ordinary regression analysis was chosen to estimate the regression model which does not consider the variation of data among different fish groups; and hierarchical Bayesian regression analysis was selected to estimate fish group-specific regression models which are calculated independently for different fish groups, taking into account the variation of data among the different groups.

3.2.2 Materials and Methods

Construction of BCF datasets

Datasets of fish BCF with metal exposure in fresh water were constructed for Cu, Zn, Cd, Pb and As independently. For the purpose of this study, As is included as a metal in convenience although it is a metalloid. Following the process by which McGeer et al. (2003) reviewed data, we did not distinguish between BCF and BAF and lumped them for the data collection because these two types of data are calculated in an identical manner and considered similarly in the regulatory context. The fish BCF database consists of different sources of data. One of the main sources is the ECOTOX Database built on the website of U.S.EPA (http://cfpub.epa.gov/ecotox). The data in the source were filtered by several search parameters (e.g. chemical, taxonomic group, endpoints, exposure medium) to customize our final ECOTOX dataset. The endpoints selected on sorting were BCF, BAF, and BCDF (dry weight basis). To

harmonize the data of BCDF into BCF and BAF, the original values of BCDF were divided by a conversion factor. The factor contains potential variation among different fish species and the variation should be treated carefully. Nevertheless, in the present study the factor 5 was chosen as a representative value according to the approach used in McGeer et al. (2003) where BCDF values were multiplied by a conversion ratio 0.2 (for fish) and also the description in Burger et al. (2003).

The other main source is the BCF dataset established by McGeer et al. (2003). They reviewed a wide range of literature, evaluating the data suitability in such a way that both exposure and whole-body metal levels should be measured by accepted analytical techniques and an exposure assessment based on standard BCF test methodologies. The rest of the BCF data come from individual papers (Balasubramanian et al. 1995; De Conto Cinier et al. 1999; Wepener et al. 2000; Coetzee et al. 2002; Velcheva 2002; Huang et al. 2003; Vinot and Pihan 2005; Arain et al. 2008; Shah et al. 2009; Vicente-Martorell et al. 2009). If the literature did not contain calculated BCF values, these were estimated mainly by the fish-body concentration divided by exposure data.

All the collected data were filtered by the response sites in the fish body where metals were accumulated. Conventionally, the chemical concentration has been expressed on a whole body basis and McGeer et al. (2003) compiled metal BCF data for a variety of aquatic organisms on this basis. For these reasons, we took the BCF data for fish whole body in the cases of Cd, Cu, and Zn. However, for Pb and As, due to scarce data for the fish whole body, those data were supplemented with the BCF data for the fish muscle which is assumed to be the main edible part for the humans and thus the most relevant for human health risk assessment (HHRA). In terms of environmental risk assessment (ERA), those two types of data need to be treated separately for their data analyses. In that sense, it can be interpreted that the analyses made for Pb and As are more relevant for HHRA rather than for ERA. In addition, the collected data were filtered by the exposure duration of the experiment. In the present study, the BCF data with more than 28 d of experimental exposure are regarded as acceptable because 28 d can be needed to reach the equilibrium of the metal concentration in fish (McGeer et al. 2003). The metal concentrations of fish sampled from the natural environment and accounted for as BAF are assumed to be at equilibrium.

The total numbers of BCF-exposure concentration data used for the analyses in the present study are 89, 40, 72, 31 and 25 for Cd, Cu, Zn, Pb and As, respectively. The variability in the collected BCF data for each metal is presented by full range (between minimum and maximum

values) and 90% interval (between the values at 5th percentile and at 95th percentile) in Table 3.4. It should be noted that certain limitations in data validity for Pb and As exist because the BCF-exposure concentration data for both fish whole body and fish muscle are combined for regression analysis under the rough assumption that there is no significant difference between the BCF – exposure correlations for fish whole body and fish muscle. The constructed dataset for each metal was divided into different fish family groups. For example, the data originally collected for the fish species such as *Salvelinus fontinalis* and *Oncorhynchus mykiss* were categorized into the same fish family group, Salmonidae. The fish species such as *Lepomis macrochirus*, *Lepomis gulosus*, and *Micropterus salmonides* were categorized together into the fish family group, Centrarchidae.

Table 3.4: Orders of magnitude (by full range and 90% interval) for the collected data of bioconcentration factor (BCF) for each metal

	Orders of magnitude (full range)	Orders of magnitude (90% interval)
Cd	4.70	3.48
Cu	1.89	1.69
Zn	3.58	2.39
Pb	3.87	3.42
As	4.66	3.74

Ordinary regression analysis

According to the findings of an inverse relationship between the collected data of BCF for fish and corresponding metal exposure (McGeer et al. 2003; DeForest et al. 2007), ordinary regression analysis was chosen to correlate BCF values with exposure concentrations in the dataset and to estimate the regression coefficients and residual error. In this analysis, all the data are jointly analyzed without considering the grouping of the data by fish families. The natural logarithms of both the BCF and the exposure concentration (hereafter ln(BCF) and ln(Cw), respectively) are set as the response variable and explanatory variable, respectively. The ordinary regression model is defined as follows,

$$\ln(\text{BCF}_i) = \alpha + \beta \ln(\text{Cw}_i) + \text{Error} \qquad (3.4)$$

$$\text{Error} \sim N(0, \text{res}_{SD})$$

where α and β are the regression coefficients (intercept and slope, respectively) of the regression line. Cw_i is the total metal concentration in water (µg/L) for data i and res_{SD} is residual standard

deviation. The linearity and normality of Equation (3.4) were analyzed by residual and normality plots.

Estimation of a generic PDF of BCF from the ordinary regression analysis

In the present study, as an example, an exposure concentration of each metal was chosen to test how much the variation contained originally in collected BCF data can be reduced when the correlation between the BCF and metal exposure is taken into account by the regression analysis. The selection of the exposure concentration was made by analyzing the data of ecotoxicological chronic NOECs (No-Observed Effect Concentrations) for several fish species. The data were collected from three databases: the ECOTOX, EAT (ECETOC aquatic toxicity: http://www.ecetoc.org/), and RIVM databases (RIVM 1999). The variability contained in the NOECs data were statistically analyzed by estimating a chronic Species Sensitivity Distribution (SSD) for each metal according to the method described in Duboudin et al. (2004) and Brock et al. (2009). In this method, a log-normal distribution was fitted to the NOECs data for each metal and confidence limits of the SSD were derived from bootstrap sampling. From the confidence limits of the SSD, the best estimate SSD (50% confidence) was chosen and the exposure concentration used in our study was taken from the hazardous concentration 5% ($HC_{5\%}$) defined in the best estimate SSD because the concentration is assumed to be low-level contamination in water and so will not strongly influence the fish behaviour and their bioaccumulation. By definition, this $HC_{5\%}$ could affect 5% of fish species, i.e. 95% of the fish species could be protected with 50% confidence. The $HC_{5\%}$ gained from the best estimate SSDs for metals are 0.55, 1.0, 17.3, 18.8 and 98.6 µg/L for Cd, Cu, Zn, Pb and As, respectively.

A generic PDF of BCF at the $HC_{5\%}$ defined for each metal was estimated based on the prediction interval of ln(BCF). The 90% prediction interval at the $HC_{5\%}$ was estimated by the equation as follows;

$$\ln(BCF) = \alpha + \beta \ln(HC_{5\%}) \pm t(0.95, n-2) \times res_{SD} \times \sqrt{\left[1 + \frac{1}{n} + \frac{(\ln(HC_{5\%}) - \ln(Cw_{all}))^2}{\sum_{i=1}^{n}(\ln(Cw_i) - \ln(Cw_{all}))^2}\right]}$$

(3.5)

where t(0.95, n - 2) denotes the 95th percentile of the Student's t-distribution with n - 2 degrees of freedom. n is the number of data points (e.g., n = 89 for Cd), and $\ln(Cw_{all})$ is the average value of all the logarithmic Cw_i i data points. The rationale to use a prediction interval estimated from Equation (3.5) rather than use the interval gained from Equation (3.4) is that we need to consider the sources of uncertainty in α, β and res_{SD} for predicting the response variable ln(BCF).

Hierarchical Bayesian regression analysis

Using the same dataset as in ordinary regression analysis, hierarchical Bayesian regression analysis was used to derive the PDFs for different fish groups. Unlike analyzing all the BCF–exposure dataset together and estimating only a set of the coefficients α and β, the hierarchical Bayesian regression approach is capable of estimating separately the sets of coefficients α_j, β_j corresponding to different fish family groups j (hereafter fish groups); moreover, it also permits them to be partially connected to each other with regard to the variability between the fish groups and the individual data size within each fish group. In this way, the coefficients for fish groups with a small number of observations can be affected to a higher extent by the data in other groups. On the other hand, the coefficients for fish groups with higher precision are dominated by their own observations. The hierarchical structure for data i within fish groups j is defined in Equation (3.6),

$$\ln\left(BCF_{i,j}\right) = \alpha_j + \beta_j \ln\left(Cw_{i,j}\right) + Error \qquad (3.6)$$

$$Error \sim N(0, res_{SD})$$

$$\alpha_j, \beta_j \mid \mu, \Sigma \sim MV(\mu, \Sigma)$$

$$\mu = \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_\alpha & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta \end{pmatrix}$$

The intercept α_j and slope β_j for fish groups j are modelled as independent outcomes of a multivariate normal population distribution (MV in Equation (3.6)) which functions as first level prior distribution thereby interconnecting the fish groups. The hyper parameter μ is defined in terms of the population intercept μ_α and slope μ_β while Σ models the variability between the groups for the intercept and slope and also the correlation between them.

To make the model in Equation (3.6) fully specified, top-level prior distributions must also be assigned to the unknown hyper parameters. Because no prior knowledge was available at this level, non-informative distributions were required. The unconstrained parameters μ_α and μ_β were assigned uniform distributions, σ_α and σ_β were assigned uniform distributions bounded below by zero, and the residual standard deviation res_{SD} a uniform distribution on log scale. The correlation ρ was assigned a uniform distribution with bounds (-1, 1). These assigned prior distributions correspond to the ones used for this model in Gelman and Hill (2007).

The posterior distribution

The joint likelihood of the unknown parameters from the fish groups conditioned on observed data is given (due to the independency of the groups and of outcomes within them) by multiplication of the normal probability density functions (f_N) of each group and of each n_j data within each group j = 1,….,J according to Equation (3.7).

$$L(\alpha, \beta, \text{res}_{SD} | \ln(Cw), \ln(BCF)) = \prod_{j=1}^{J} \prod_{i=1}^{n_j} f_N(\ln(BCF_{i,j}) | \alpha_j, \beta_j, \text{res}_{SD}, \ln(Cw_{i,j})) \quad (3.7)$$

where L stands for the likelihood function. The joint population distribution of the coefficients α_j, β_j, $j = 1,\ldots,J$ is given by multiplication of J multivariate normal distribution functions according to Equation (3.8).

$$f(\alpha, \beta | \mu, \Sigma) = \prod_{j=1}^{J} f_{MV}(\alpha_j, \beta_j | \mu, \Sigma) \quad (3.8)$$

The joint posterior is given by multiplication of the joint likelihood in Equation (3.7), the population distribution in Equation (3.8) and the prior distributions of res_{SD}, μ and Σ.

$$f(\alpha, \beta, \text{res}_{SD}, \mu, \Sigma | \ln(Cw), \ln(BCF)) \propto L(\alpha, \beta, \text{res}_{SD} | \ln(Cw), \ln(BCF)) \times f(\alpha, \beta | \mu, \Sigma) \times$$
$$f(\text{res}_{SD}) \times f(\mu) \times f(\Sigma) \quad (3.9)$$

The WinBUGS software package (freely available at http://www.mrc-bsu.com.ac.uk/bugs/) was used to draw inferences on unknown parameters in the model in Equation (3.6) by sampling from the joint posterior in Equation (3.9). WinBUGS requires proper prior distributions (i.e. distributions with finite integrals) and for this reason finite bounds were used for the uniform prior distributions of the top level parameters. The bounds were chosen to be large enough to include the probable values of the parameters. To assure that these arbitrary bounds did not seriously influence the results, the bounds were also assessed and made even wider if needed after obtaining the results.

### *Estimation of species-specific PDFs of BCF from hierarchical Bayesian regression analysis*

Species-specific PDFs of BCF were estimated at the $HC_{5\%}$ defined previously. Fish group-specific PDFs of BCF at the $HC_{5\%}$ can be estimated by the following process. Ten thousand samples of α_j and β_j are drawn randomly from their posterior distributions. The distribution of $\ln(BCF)$ at $HC_{5\%}$ was simulated by drawing each value of $\ln(BCF_j^k)$, $k = 1, \ldots 10{,}000$ from the normal distribution $N(\alpha_j^k + \beta_j^k \ln(HC_{5\%}), \text{res}_{SD}^k)$ where each $\alpha_j^k, \beta_j^k, \text{res}_{SD}^k$ is taken as k_{th} sampled value taken from the posterior distribution.

3.2.3 Results and Discussion

Ordinary regression analysis

Table 2 presents 90% confidence intervals of coefficients α, β gained from the ordinary regression analysis of the full dataset for each metal together with the residual standard deviation (res_{SD}) and the coefficient of determination R^2. Figure 3.2 visualizes the regression line made for all the data points with the boundaries of 90% prediction interval. In Table 3.5, the highest R^2 seen for Zn (0.85) indicates that the regression model for Zn have the potential to estimate more accurately the BCF at a certain exposure concentration, compared with the

regression models for other metals. In the next section, the regression models for metals are compared at given exposure concentrations ($HC_{5\%}$).
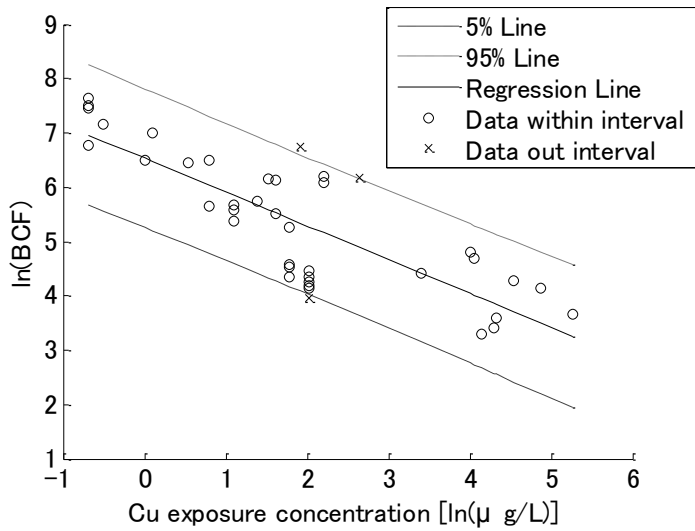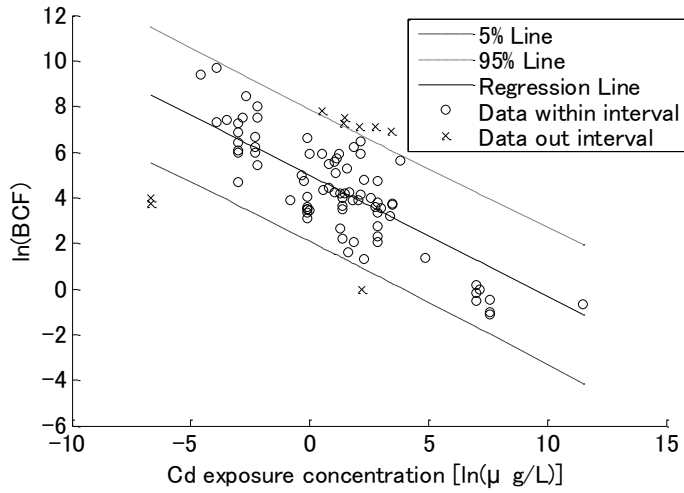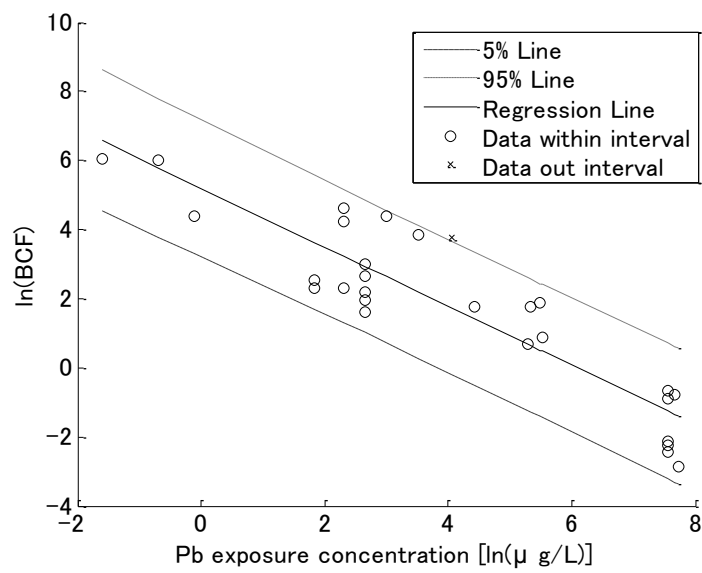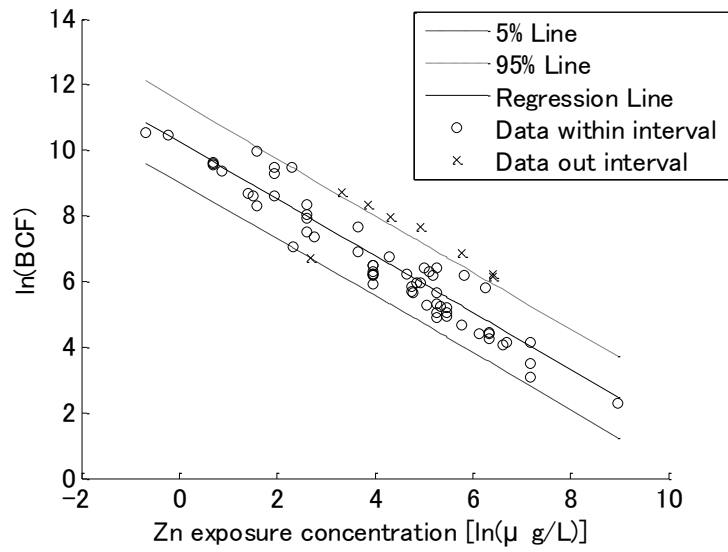
Figure 3.2: The regression line made for all the dataset with the boundaries of 90% prediction interval for each metal

Table 3.5: 90% intervals and means of coefficients α and β, residual standard deviation (res_{SD}) and the coefficient of determination (R^2) for each metal

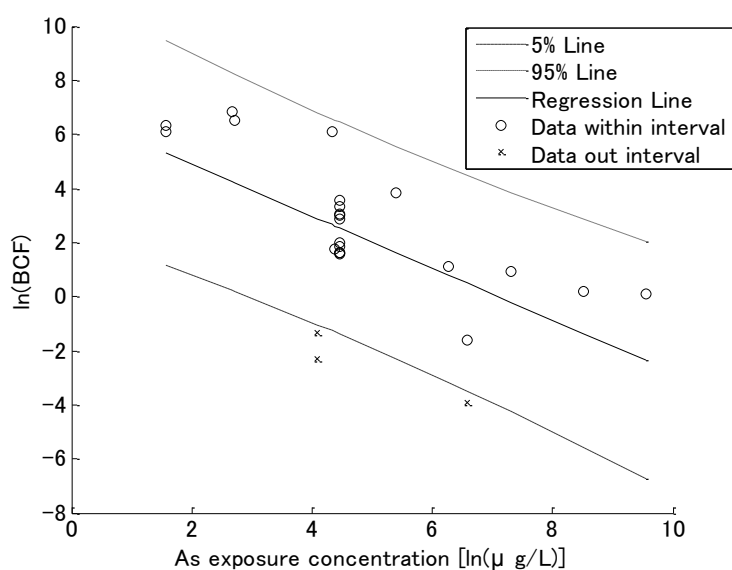|	α			β				
	5%	mean	95%	5%	mean	95%	res_{SD}	R^2
Cd	4.66	4.99	5.31	-0.62	-0.53	-0.43	1.73	0.49
Cu	6.23	6.53	6.84	-0.74	-0.62	-0.50	0.74	0.66
Zn	9.92	10.25	10.59	-0.94	-0.87	-0.79	0.72	0.85
Pb	4.58	5.20	5.82	-0.98	-0.85	-0.73	1.11	0.81
As	4.67	6.83	8.98	-1.38	-0.96	-0.54	2.24	0.40

Estimation of a generic PDF of BCF from the ordinary regression analysis

The 90% prediction interval of ln(BCF) at the $HC_{5\%}$ was transformed into that of a normal-scaled BCF (Table 3.6) with the orders of magnitude over the interval and the geometric mean (GM) and geometric standard deviation (GSD) which represent the PDF of BCF. The smallest GSD and 90% interval are seen in the BCF distribution for Zn compared to the GSD for other metals, which reflects the highest accuracy of the regression model estimated for Zn. Compared with the orders of magnitude (90% interval) that the data of BCF range over, those of 90% prediction interval of BCF have the narrower range at the given exposure concentration. The relative differences between them for Cd, Cu, Zn, Pb and As are 0.28, 0.34, 0.56, 0.51, and 0.09, respectively. In the present study, the relative difference was calculated as ($data_{90}$ − $pred_{90}$) / $data_{90}$, where $data_{90}$ and $pred_{90}$ indicate 90% interval of data variation and 90% prediction

74

interval, respectively. This indicates that, for most of the metals (Cd, Cu, Zn, Pb), when a specific exposure concentration of metal is known, the PDFs of the current and future observations of BCF can be estimated more informatively, i.e., less uncertainly by using the regression models than by directly analyzing BCF data regardless of the corresponding exposure concentrations. Such PDFs with less uncertainty can be useful for the probability analysis conducted in the modelling studies dedicated to environmental and health risk assessment since they can provide accurate model predictions for the potential risk of metals.

Table 3.6: 90% prediction interval (with orders of magnitude for the interval) and generic probability density function (PDF) of bioconcentration factor (BCF) at the hazardous concentration 5% (HC5$_\%$) for each metal

	HC$_{5\%}$(μg/L)	90% Interval of BCF		Orders of magnitude	Generic PDF of BCF (log normal)	
		5%	95%		GM[b]	GSD[b]
Cd	0.55	11.1	3627	2.51 (3.48)[a]	201	5.7
Cu	1	192	2462	1.11 (1.69)	687	2.1
Zn	17.3	716	8103	1.05 (2.39)	2408	2.1
Pb	18.8	2.2	101	1.67 (3.42)	14.9	3.1
As	98.6	0.2	559	3.40 (3.74)	11.2	9.8

GM and GSD denote geometric mean and geometric standard deviation, respectively.

[a] The values in bracket are the orders of magnitude (90% interval) for the collected BCF data.

[b] GM and GSD were calculated as;

$$GM = \exp\big(\alpha + \beta \ln(HC_{5\%})\big), GSD = \exp(S)$$

$$S = \text{res}_{SD} \times \sqrt{\left[1 + \frac{1}{n} + \frac{(\ln(HC_{5\%}) - \ln(Cw_{all}))^2}{\sum_{i=1}^{n}(\ln(Cw_i) - \ln(Cw_{all}))^2}\right]}$$

Hierarchical Bayesian regression analysis

The posterior estimates of coefficients α_j, β_j, $j = 1,\ldots,J$ for each metal are presented in the form of 90% intervals (Figure 3.3). Table 3.7 shows the names of fish groups corresponding to the numbers on the horizontal axis in Figure 3.3 and the number of data points (set of BCF- metal exposure data) for each fish family. For each metal, the posterior 90% intervals of coefficients α_j and β_j (Figure 3.3) are found to be much wider than the common intervals of the coefficients α and β (Table 3.5). For example, the average posterior 90% intervals of coefficients α_j and β_j over J for Cu are respectively more than 3 and 4 times as wide as the common interval of the coefficients α and β. This significant gap between these intervals is because the estimation by

hierarchical Bayesian regression analysis takes into account the variability between the coefficients of different fish groups.

Table 3.7: Names of fish families and their data points

N	Name of fish family / data points				
	Cd	Cu	Zn	Pb	As
1	Anguillidae / 1	Anguillidae / 1	Atherinopsidae / 1	Anguillidae / 1	Bagridae / 2
2	Atherinopsidae / 1	Atherinopsidae / 1	Catostomidae / 2	Catostomidae / 1	Cichlidae / 6
3	Balitoridae / 1	Catostomidae / 1	Centrarchidae / 9	Centrarchidae / 3	Cyprinidae / 11
4	Catostomidae / 2	Centrarchidae / 4	Clariidae / 3	Channidae / 1	Percichthyidae / 4
5	Centrarchidae / 26	Clupeidae / 1	Clupeidae / 1	Cichlidae / 2	Soleidae / 1
6	Cichlidae / 3	Cyprinidae / 2	Cyprinidae / 3	Clariidae / 3	Sparidae / 1
7	Clariidae / 3	Esocidae / 2	Cyprinodontidae / 1	Cyprinidae / 9	
8	Clupeidae / 1	Gasterosteidae / 3	Esocidae / 1	Esocidae / 1	
9	Cyprinidae / 3	Percidae / 2	Fundulidae / 3	Osteichthyes / 3	
10	Cyprinodontidae / 13	Salmonidae/ 23	Percidae / 2	Percidae / 2	
11	Esocidae / 1		Poeciliidae / 6	Salmonidae / 3	
12	Gasterosteidae / 1		Salmonidae / 30	Soleidae / 1	
13	Percidae / 2			Sparidae / 1	
14	Poeciliidae / 3				
15	Salmonidae / 28				

It is shown in this analysis that the coefficients for the fish groups with rich information, i.e., more than 10 data points, have narrower posterior intervals than those with poor information, i.e., only 1 or 2 data points. For example, in the case of Zn, the ratio of average 90% interval of β for the groups with 10 or more data points to the average intervals for the groups with 1 or 2 data points is 41 %. The ratios for Cd, Cu, Pb and As are 77, 12, 54 and 31%, respectively. For Pb, the interval for the group with 9 data points was compared with the average interval for the rest of groups with less than 9 data points. These results track the description above about how hierarchical Bayesian regression analysis estimates the coefficients for the groups with a poor dataset and for those with a rich dataset with high precision.
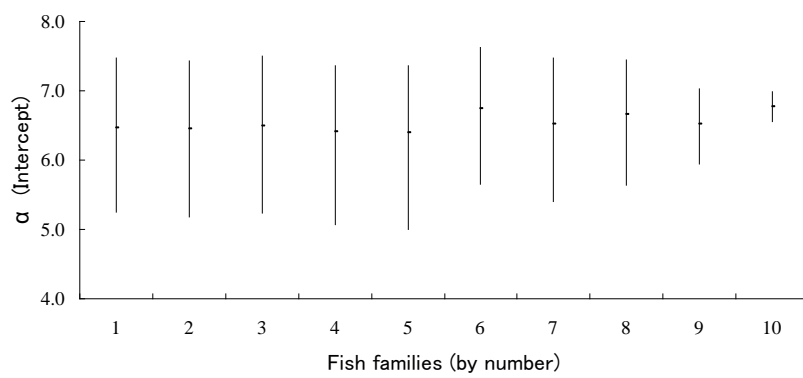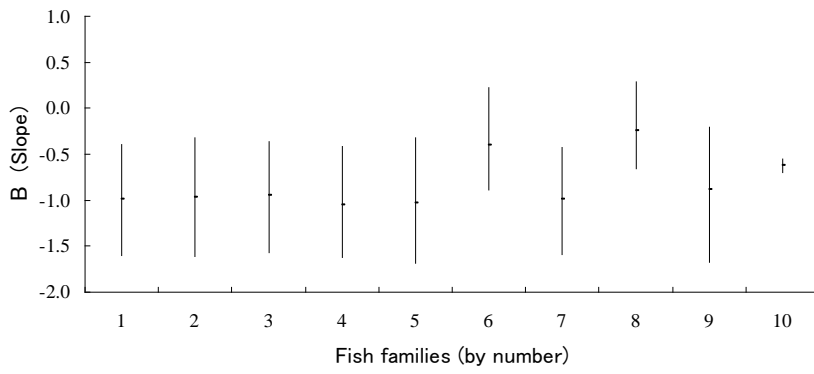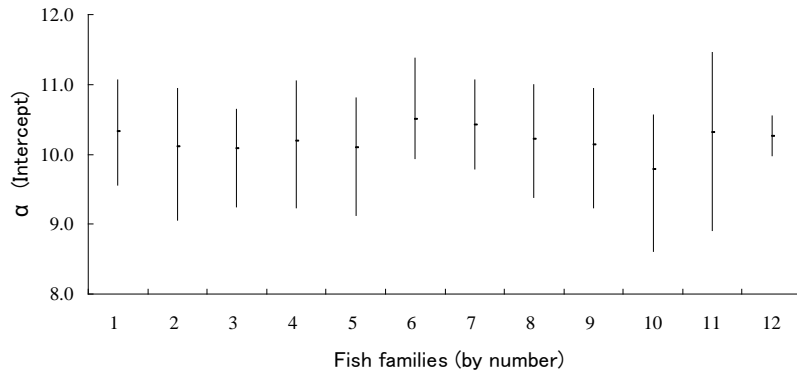
Cd
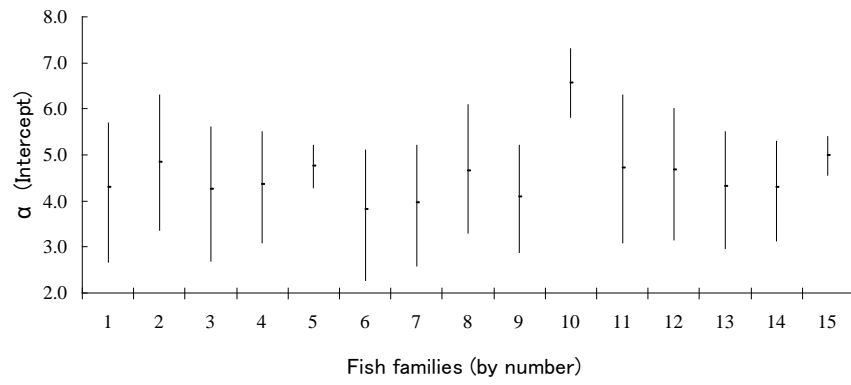
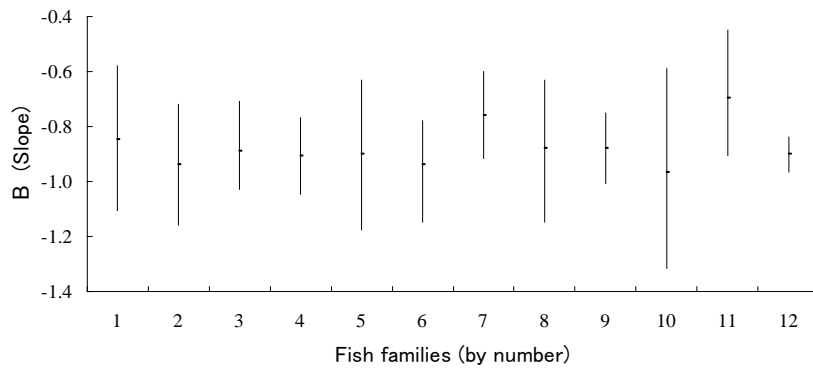

Cd
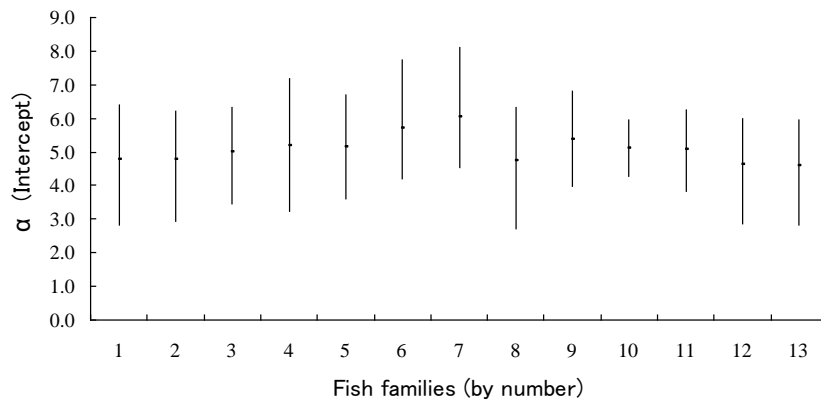

Cu

Cu

Zn
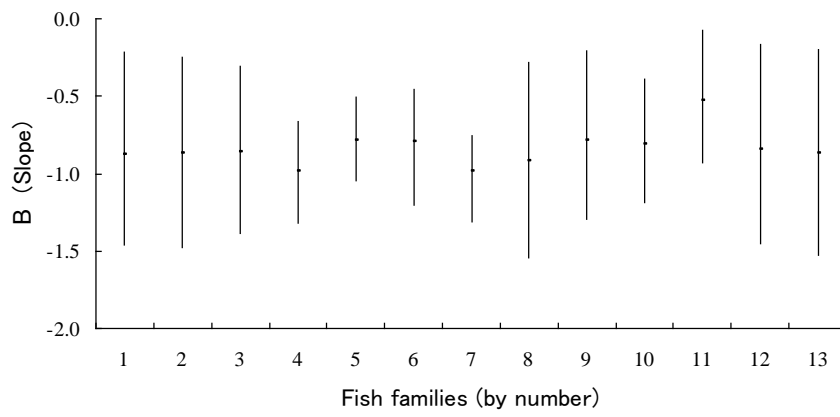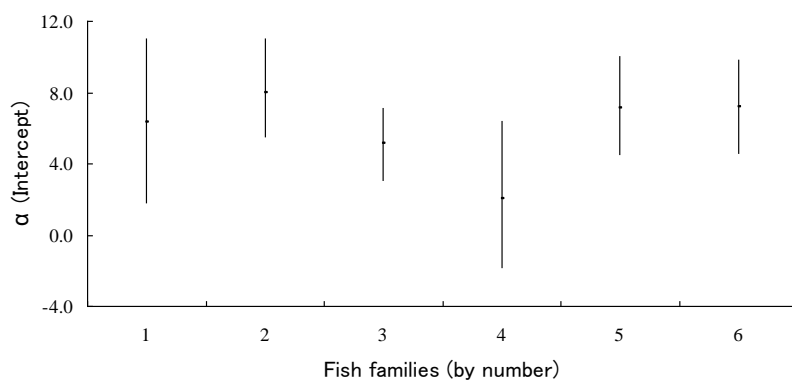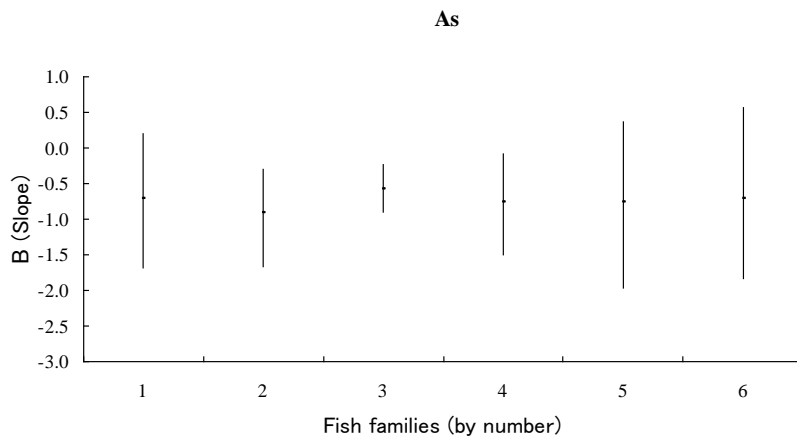

Zn

Pb

Pb

As
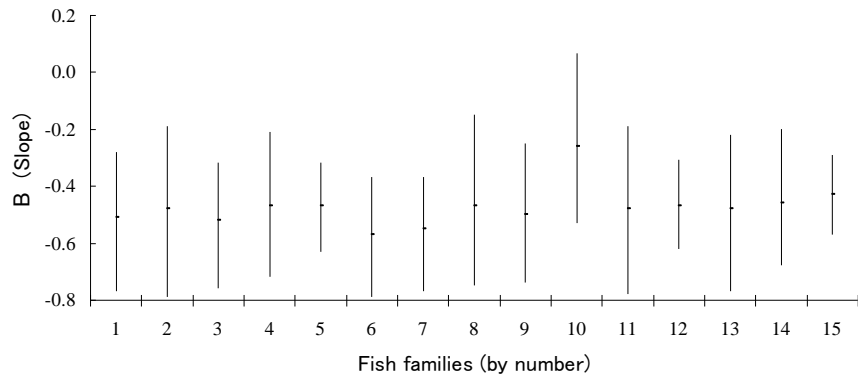
As

Figure 3.3: Posterior 90% intervals and mean values of coefficients α_j and β_j (j denotes the number of fish family) for each metal. The points in the middle of interval lines indicate mean values. The top and bottom in the lines correspond to the values at 95th percentile and 5th percentile, respectively. The numbers at horizontal axis in the figure correspond to the numbers given to fish families in Table 3.7.

Estimation of group-specific PDFs of BCF from hierarchical Bayesian regression analysis

Fish group-specific PDFs of BCF at the $HC_{5\%}$ were simulated by the process discussed in the previous section. Table 3.8 shows the fish group-specific PDFs of BCF for metals with the 90% posterior intervals. As seen in the difference of coefficient intervals estimated among fish groups, it is found for each metal that the 90 % posterior intervals of BCF estimated for the fish groups with more information tend to have narrower ranges than those with less information. The average orders of magnitude over the 90% posterior intervals of BCF for different fish groups are 2.37, 1.09, 0.93, 1.67 and 2.75 for Cd, Cu, Zn, Pb and As, respectively. The relative differences from the 90% interval of BCF data range for Cd, Cu, Zn, Pb and As are 0.32, 0.35, 0.61, 0.51 and 0.27, respectively. The estimated average 90% posterior intervals are found to be almost equal to or narrower than those derived from ordinary regression analysis (Table 3.6) although the posterior intervals of the coefficients are found to be much wider than those calculated by ordinary regression analysis. The main reason for this lower uncertainty in the posterior BCF intervals is that, in this simulation for each metal, estimated values of the residual standard deviation (res_{SD}) in Equation (3.6) that accounts for the within-fish-group variation were lower than the residual standard deviation given in Equation (3.4) which represents the high variation contained in all the data. When looking into the PDF estimated for each fish group, the PDFs estimated for some fish groups were found to have low uncertainty, and thus can be useful for the probability analysis in order to make accurate model predictions for the potential risk of metals. For example, in the cases of Cu and Zn, the orders of magnitude in the 90% intervals are less than one for two fish groups and ten fish groups, respectively.

Table 3.8: 90% prediction interval (with orders of magnitude for the interval) and fish group-specific probability density function (PDF) of bioconcentration factor (BCF) at the hazardous concentration 5% ($HC_{5\%}$) for each metal

(1) Cd

N	Data points	90% Interval of BCF 5%	90% Interval of BCF 95%	Orders of magnitude	Group-specific PDF of BCF (log normal) GM	Group-specific PDF of BCF (log normal) GSD
1	1	5.8	1808	2.50	95.6	5.5
2	1	10.4	2981	2.46	172	5.5
3	1	6.1	1808	2.47	97.5	5.5
4	2	6.4	2208	2.54	107	5.5
5	26	14.4	1636	2.05	159	4.5
6	3	2.9	1097	2.58	62.2	6.0
7	3	4.4	1212	2.44	73.7	5.5
8	1	8.0	1636	2.31	136	5.0
9	3	6.1	1097	2.25	84.8	5.0
10	13	65	12088	2.27	796	5.0
11	1	9.5	2697	2.45	155	5.5
12	1	8.5	2208	2.41	144	5.5
13	2	7.2	1480	2.31	103	5.0
14	3	5.6	1480	2.42	92.8	5.5
15	28	16.0	1998	2.10	183	4.5
*	89	11.1	3627	2.51	201	5.7

GM and GSD denote geometric mean and geometric standard deviation, respectively.

*The values in the last row present the total data poinnts, the 90% prediction interval of BCF (with orders of magnitude for the interval), and the generic PDF of BCF being estimated by ordinary regression analysis (see Table 3.6).

(2) Cu

N	Data points	90% Interval of BCF 5%	90% Interval of BCF 95%	Orders of magnitude	Group-specific PDF of BCF (log normal) GM	Group-specific PDF of BCF (log normal) GSD
1	1	156	2303	1.17	639	2.5
2	1	158	2294	1.16	626	2.7

N	Data points	5%	95%	Orders of magnitude	GM	GSD
3	1	145	2319	1.20	659	2.6
4	4	126	2169	1.23	596	2.6
5	1	109	2156	1.30	578	2.7
6	2	237	2768	1.07	854	2.2
7	2	161	2500	1.19	665	2.6
8	3	204	2340	1.06	742	2.1
9	2	255	1760	0.84	652	1.8
10	23	412	1953	0.68	863	1.6
*	40	192	2462	1.11	687	2.1

(3) Zn

N	Data points	90% Interval of BCF		Orders of magnitude	Group-specific PDF of BCF (log normal)	
		5%	95%		GM	GSD
1	1	863	8185	0.98	2670	2.0
2	2	528	4770	0.96	1652	1.9
3	9	679	4770	0.85	1863	1.8
4	3	665	6063	0.96	1978	2.0
5	1	596	5541	0.97	1863	2.0
6	3	880	7187	0.91	2566	1.9
7	11	1525	9897	0.81	3790	1.8
8	1	685	7115	1.02	2231	2.1
9	3	742	5943	0.90	2080	1.9
10	2	361	3429	0.98	1108	2.0
11	6	1164	11731	1.00	3866	2.0
12	30	889	5943	0.83	2231	1.8
*	72	716	8103	1.05	2408	2.1

(4) Pb

N	Data points	90% Interval of BCF		Orders of magnitude	Group-specific PDF of BCF (log normal)	
		5%	95%		GM	GSD
1	1	1.3	55.3	1.64	8.7	3.2
2	1	1.2	67.4	1.76	8.9	3.5

3	3	2.1	65.9	1.50	11.5	2.8
4	1	1.2	71.8	1.78	9.7	3.5
5	2	3.1	99.7	1.51	17.5	2.9
6	3	4.4	176	1.61	28.2	3.0
7	9	3.6	155	1.63	24	3.2
8	1	1.1	50.9	1.68	7.5	3.4
9	3	2.8	118	1.62	19.5	3.1
10	2	1.7	144	1.92	15.3	3.7
11	3	5.2	202	1.59	33.4	3.0
12	1	1.1	60.9	1.73	8.2	3.2
13	1	0.9	59	1.80	7.8	3.4
*	31	2.2	101	1.67	14.9	3.1

(5) As

		90% Interval of BCF		Orders of magnitude	Group-specific PDF of BCF (log normal)	
N	Data points	5%	95%		GM	GSD
1	2	1.3	270	2.32	20.1	5.0
2	6	4.6	545	2.07	49.4	4.1
3	11	1.1	99.5	1.97	11	4.1
4	4	0.02	3.0	2.17	0.2	4.5
5	1	0.3	2981	3.98	36.6	22.2
6	1	0.7	6634	3.95	49.4	22.2
*	25	0.2	559	3.40	11.2	9.8

3.2.4 Conclusion

The generic PDF for fish BCF given by ordinary regression analysis for each metal does not present detailed information of how the BCF distribution corresponding to individual fish families is shown differently from others. However, this general information can be useful, for example, when it is required as an input parameter in multi-media modelling studies where the real environmental system with much complexity is often considerably simplified and the simplified model may require generalized input information rather than a detailed or specific one.

On the other hand, fish group-specific PDF estimated by hierarchical Bayesian regression analysis can tell us more exhaustive information considering the differences residing among different fish groups, and thus this information can be helpful, for instance, when environmental modelling studies deal with the site-specific (local) scenario where an individual trend on BCF for the site-specific fish group may be necessary. A remarkable feature of this analysis is that it is capable of the estimation of BCF range and distribution even for a fish group with scarce information, i.e., 1 or 2 data points by partially relating it to the data for other fish groups.

The present study shows that in both ordinary regression and hierarchical Bayesian regression analyses, the estimated PDF and range of BCF for each metal have less uncertainty than the variation of collected BCF data, when an exposure concentration is given. This signifies that the BCF ranges and PDFs estimated for metals by both statistical approaches can tell us more informative knowledge about current and future BCF observations, correlating them to the exposure concentrations, and thus use of such PDFs for the probability analysis conducted in the modelling studies dedicated to environmental and health risk assessment will lead to the accurate predictions of the potential risk concerning metals.

Two databases were used for the data collection without deeply reviewing each of the references listed on the databases because of the limitation of the information we could access at the moment and also combined the data of BCF and BAF for the data analyses; thus, a certain vagueness may occur in the validity of collected data and estimated regression models. In the present study, we assume such vagueness is considered in the uncertainty the estimated regression models contain. Nevertheless, in the future study, in order to improve the quality of data to analyze and regression models, it may be important to conduct a deep review of each source in the databases and also to analyze separately the data of BCF and BAF.

3.3 *Probabilistic parameterization for the other input parameters used in freshwater compartment*

In freshwater compartment, the partition of contaminants between water and suspended particulate matter (SPM) can be considered as one of main transfer factors for chemical pollutants. The ratio of neutral organic compound in SPM to that in water is expressed by the distribution coefficient $K_{d,SPM}$ (m^3/g). The $K_{d,\,SPM}$ is defined by the substance-dependent organic carbon partition coefficient K_{OC} and the time-dependent fraction of organic carbon in SPM, $y_{OC,SPM}(t)$:

$$K_{d,SPM}(t) = 10^{-6} \cdot y_{OC,SPM}(t) \cdot K_{OC} \qquad (3.10)$$

In many cases, a point value (i.e., mean) of riverine organic fraction in SPM ($y_{OC,SPM}$) is used for modelling studies. Since fluvial dynamics are characterized by significant spatial and temporal variations especially caused by rainy events, their influence on the organic carbon fraction in SPM ($y_{OC,SPM}$) should be considered. In the framework of a project called SCOPE (Ittekot and Lane 1991), riverine particulate organic carbon (POC) was monitored for major rivers located all around the world, and the observed relations between POC and SPM for these rivers were found to be similar. The variation of SPM concentration between 0–50 mg/L corresponds to low water discharge, whereas the SPM variation between 50–300 mg/L is often associated with the intermediate water discharge and resuspension processes of materials into the river bed. The SPM concentration greater than 300 mg/L corresponds to the intensive soil erosion caused by running water flow and by the increase of the number of flood events. These different origins of SPM lead to the fact that for low SPM concentrations, the organic fraction in SPM tends to be high (3-25 % of SPM) and is essentially phytoplanktonic. For higher SPM concentrations, the organic fraction tend to be low (2-3 % of SPM) and is essentially allochthonous (Veyssy et al. 2004). According to results reported in the SCOPE project, it was thus proposed to use the following function to correlate $y_{OC,\ SPM}(t)$ and total SPM (g/m^3) :

$$y_{OC,SPM}(t) = \frac{\alpha}{SPM(t)} + \beta \qquad (3.11)$$

where α and β are empirical calibration parameters. The probability density function (PDF) for α and β were derived from the values proposed in Ittekot and Lane (1991) (minimum, mean and maximal values of the fraction of POC for each range of SPM level).

The time-dependence of the Suspended Particulate Matter SPM(t) can be obtained from the empirical relationship called sediment rating curve (Asselman 2000; Syvitski et al. 2000; Morehead et al. 2003):

$$SPM(t) = a \cdot Q(t)^b \qquad (3.12)$$

where a and b are empirical calibration parameters and Q(t) is time-dependent flow rate (m^3/s). In this section, the approach to derive the PDF of the empirical calibration parameters, a and b is described. The empirical calibration parameters, a and b in Equation (3.12) were estimated using a Bayesian approach. The Bayesian approach used for the estimation of these parameters is elaborated in Ciffroy et al (2010b). As a starting point, the database set by Syvitski et al. (2000) was used for this approach. The database contains the values of $\log_{10}(a)$ and b for 57 rivers which were collected by long-term monitoring. For each of these rivers, Syvitski et al. (2000) provided best estimates for $\log_{10}(a)$ and b, as well as geographical and physical characteristics of

the river watersheds (latitude, longitude, mean annual discharge, minimum and maximum temperatures, peak flow, and etc). The study also tested with the database several relationships to relate $\log_{10}(a)$ and b to some of these characteristics, and found, for instance, that latitude and mean annual discharge are reliable predictors of $\log_{10}(a)$ and b. Based on this knowledge, the PDFs of $\log_{10}(a)$ and b were defined by using the information of the Seine watershed characteristics. In the Bayesian approach of this study, the PDFs were used as prior distributions and were updated by the actual monitoring data, i.e., monthly data of SPM and flow rates measured at the same date during the period 1985-1995, which was obtained from the database of Seine-Normandy Water Management Agency (SNWMA).

The dataset was subdivided into two part: 10% of data were used to update the log(a) and b distributions by the Bayesian approach, and 90% of data were used as a validation dataset for the SPM estimation made based on the posterior distributions of log(a) and b. The estimated posterior distributions for $\log_{10}(a)$ and b are N(-4.19, 0.33) and N(0.99, 0.13), respectively. Based on Equation (3.12), the distribution of SPM was estimated by generating random values from these posterior distributions. It was observed that there is a good agreement between mean values of the estimated distribution and the validation dataset with a bias (median of relative errors) of 31%. Almost all the SPM values in the validation dataset were found to be contained in the 90% confidence interval of the estimated SPM distribution.

The other approaches to estimate PDFs for the other input parameters used in freshwater compartment are presented in a deliverable in 2-FUN project.

4　Application of 2-FUN tool based on a case study

4.1　*Introduction*

The paradigm of health risk assessment may consist of two main pillars, i.e., the exposure and dose-response assessments. Exposure assessment frequently involves the process of estimating or measuring the magnitude, frequency and duration of exposure to chemicals, along with the number and characteristics of the population exposed. Human exposure to chemicals via multiple pathways such as drinking water, inhaled air, foods, and etc) can be estimated by the multimedia models that calculate the distribution of chemicals in the component media, i.e., air, water, soil, plants, cattle, and human media (e.g., SimpleBox (Brandes et al. 1996), QWASI (Mackay 1983; Di Guardo et al. 2006; Warren et al. 2007), Trim.Fate (USEPA 2002), and etc). Combined with the information about human behaviors such as dietary habits, time spent

outside, and etc, the multimedia models can provide an estimation of the daily chemical intake by inhalation or ingestion by the population of interest.

Once the exposure scenario is identified, the dose-response assessment is typically achieved by comparing exposure outputs (e.g., the daily intake) to the reference doses estimated from toxicological data. Currently, risk assessment methodologies involve simple dose-response models that link the external dose (e.g., environmental concentration) to the adverse effects. However, the use of such simple dose-response models does not reflect the current understanding of the mode of action of a chemical and does not facilitate extrapolations to other scenarios (species, exposure routes...). Thus the determination of internal effective concentrations, i.e., in the target tissues where toxic effects arise, is required to characterize accurately the link between an external exposure and the internal dosimetry that may be associated with the observed effects (Andersen and Dennison 2002). Physiologically based pharmacokinetic (PBPK) models have been developed to predict the internal effective concentrations, i.e., in the target internal tissues where toxic effects arise (d'Yvoire et al. 2007). Moreover, these models are well-suited for integrating available information on age- or gender-dependent changes and then evaluating the influence of these changes on the internal dosimetry (Clewell et al. 2004; Beaudouin et al. 2010). A dose-response model is then applied to link the effective concentration to the adverse effects.

Coupling of a multi-media model for different exposure pathways with a generic PBPK model for the human population allows assessing directly the impact of the exposure scenarios on the chemical's concentration in the target tissues. To our best knowledge, environmental multimedia and PBPK models have never been systematically integrated in a common platform. These models are usually developed on different software, which makes it difficult to perform the integrated analysis coupling both of models. In the framework of the European project called 2-FUN project (Full-chain and UNcertainty Approaches for Assessing Health Risks in FUture ENvironmental Scenarios), an integrated 'multimedia-PBPK' model (called 2-FUN tool) was developed. The integrated modelling tool was also designed for performing uncertainty and inter-individual variability analyses using Monte Carlo simulations, and several kinds of sensitivity analyses by several regression and Fourier approaches. The 2-FUN tool is capable of conducting lifetime risk assessments for different human populations (general population, children at different ages, pregnant women, and etc) taking into account the exposure to chemicals via multiple pathways such as drinking of water, inhalation of air, and ingestion of foods.

This kind of modelling tool is specially designed for compounds that are ubiquitous in our environment in order to identify the contribution of the different sources (e.g., food, air or drinking water) to the total intake and their contributions to the internal dosimetry. In this study, a special attention is paid to Polycyclic Aromatic Hydrocarbons (PAHs). PAHs are products of incomplete combustion of fossil fuels, wood, and other organic materials at temperatures between 300 and 600 °C. PAHs exert a variety of toxic responses toward human health, including carcinogenic, immunotoxic, neurodevelopmental and cardiovascular effects, and are likely to contribute to smoking-related diseases (Sorensen et al. 2003; Slotkin and Seidler 2009). Their carcinogenic activity is typically exerted after the metabolic activation to reactive intermediates that can damage DNA through adduct formation. Major sites of toxicity are the sites of absorption. For example, forestomach tumors have been observed in rodents after ingestion. Other sites of toxicity can be lungs in case of inhalation, skin for dermal exposure, and the sites of metabolism (e.g., the liver). Constituents of blood (e.g., white blood cells) are also of interest to detect early markers of toxicity. Benzo(a)pyrene (B(a)P), a five ring PAH, is usually used as a marker of effects of complex mixtures including PAHs, either directly (as a surrogate for the PAH fraction of complex mixtures) or using toxicity equivalency factors to relate the toxicity of each PAH to the one of B(a)P (Pufulete et al. 2004).

The aims of this study were then to perform an integrated modelling approach by the newly developed 2-FUN tool to assess the human exposure to a toxic chemical and the potential health risks due to the exposure; and to identify the input parameters and exposure pathways that are sensitive to specified model outputs, i.e., chemical concentration and accumulation in target tissues (liver and lungs), integrating uncertainty and sensitivity analyses. A case study was designed for a region situated on the Seine River watershed, downstream of the Paris megacity and for B(a)P emitted from industrial zones in the region.

4.2 *Materials and methods*

Case study

As a case study, a region situated on the Seine River watershed, just downstream of the Paris megacity was selected. The region is characterized by strong industrialization and urbanization, with industries and domestic anthropogenic activities potentially releasing toxicants. B(a)P was selected as a target chemical. This case study considers the chemical exposure to humans via the ingestion of water, fish, vegetables (root, potato, and leaf), grain, fruit, meat (beef), milk, and the inhalation of out-door air. The drinking water is assumed to be sourced from river water in this scenario.

In this case study, there are two model inputs associated with B(a)P levels; one is B(a)P concentration in river water at the upstream zone, and the other is B(a)P concentration in the air flow entering the target region. Temporal dynamics is retained in this modelling approach to allow an assessment of the impact of temporal variations of these model inputs. The data of B(a)P concentration in the upstream zone were obtained from the database made by the Seine-Normandy Water Management Agency (SNWMA) that is the organization in charge of environmental monitoring in this region. Database of the monitoring station at Poissy (the place located in the upstream about 20 km from Paris and in the downstream of the The Achères wastewater treatment plant) includes annual measurements of B(a)P concentration in bottom sediments (from 1993 to 2008, except for 1995 and 2002) and monthly measurements of B(a)P concentration in raw river water (in 2007 and 2008). The magnitude of B(a)P levels in the air flow in Paris region was found in Quéguiner et al. (2010). Long-term time series of B(a)P concentrations in the upstream and the air flow were made up based on those two kinds of information sources. The methods to make up them are described in the following section.

Other main model inputs include flow rates of the Seine River (m^3/s) and meteorological data such as air temperature (°C), precipitation (m/day), and wind velocity (m/day). Monthly flow rates were selected from the measurements monitored by SNWMA during the period of 1985-1995. Monthly datasets of air temperature and wind velocity were selected from the statistics based on observations taken between 7/2002 - 9/2010 daily from 7am to 7pm at Paris-CDG (available at http://www.windfinder.com/windstats). Daily dataset of precipitation was chosen from annual observations (monitored in 1999) at the meteorological station located in Orléans (about 120 km from the center of Paris).

Generally, the raw-river water is treated by water purification processes before it is supplied to houses as tap water, and thus removal of certain amount of B(a)P through water purification processes must be considered. The raw-river water used for the tap water passes though several treatment processes (e.g. clarification, disinfection, activated carbon filtration). Stackelberg et al (2004 and 2007) observed the efficiency of conventional water treatment processes for the removal of contaminants such as PAHs. Measurements on Anthracene, Fluoranthene, Phenanthrene and Pyrene indicated that the processes of clarification and disinfection are poorly efficient for removing PAHs, while activated carbon filtration is highly efficient. The studies estimated the removal rates for the raw water including both its dissolved and particulate phases, and thus it is difficult to estimate the removal efficiency for the dissolved phase alone. Therefore, as a worst scenario, this case study simply assumes that the particulate fraction of

B(a)P in raw-river water is completely removed by the water treatment processes, while the dissolved fraction is not.

Model outputs of interest chosen for this case study are B(a)P concentrations in liver and lungs, and also accumulated quantities of metabolites formed in liver and lungs.

Reconstruction of long-term time series of B(a)P levels in river and air

As described in the previous section, the case study requires long-term time series of B(a)P concentrations in the upstream river water, and were made up using the limited measurements of B(a)P concentrations in bottom sediments and raw river water. The annual monitoring data of B(a)P concentration in bottom sediments from 1993 to 2008 (except for 1995 and 2002) show that there is no clear time-evolution of B(a)P levels over the period. Thus, from the observation in bottom sediments, it can be assumed that the B(a)P levels in raw river water does not have the trend of their increase or decrease over the period. In this study, 23 data values of B(a)P concentration in river water, which were monitored in 2007 and 2008, were used to reconstruct daily B(a)P concentration in river water over the period from 1993 to 2008.

However, a significant fraction (56%) among these 23 available data of the concentration in river water was found to be below the limit of detection (LOD). In order to give a realistic range of B(a)P levels in river water, it is dispensable that the range of non-detects under LOD is provided with actual values. Up to now several approaches have been proposed to address this issue. The most common (and easy-to-use) approach may be simply ignoring the part of non-detection or giving substitutes for the non-detect part. The substitutes can be selected from zero, LOD itself, LOD/2 or LOD/$\sqrt{2}$ (Helsel 2006). However, such approaches can cause over- or underestimation of the mean of data distribution, and inversely under- or overestimation of the standard deviation. For example, USEPA provided a critical guidance of the reliable use of such methods in Data Quality Assessment (USEPA 2000) and suggested that substitute methods are useful if the fraction of non-detect data contained in a total dataset is fewer than 15 %. Instead of using the approaches of substitution, the present study focused on distributional methods (Baccarelli et al. 2005). The methods require defining an attribution of data distribution, e.g., Normal, Log-normal, Gamma or Gumbel distribution (Zhang et al 2004), and the parameters of the selected distribution (e.g., a set of mean and standard deviation for a normal distribution) are obtained from the maximum-likelihood estimation (Cohen 1961 and 1991). Such distributional methods can also be extended to reconstruct a complete dataset considering non-detect values. This study selected the Cohen's approach (described in detail by Zhang et al (2004) and Kuttatharmmakul et al (2001)), assuming that contamination level in raw river water follows a normal distribution. The mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ of the 'true'

distribution including non-detect values were calculated based on the mean and variance of the detected values, using a correction factor:

$$\hat{\mu} = \bar{y} - \lambda(\bar{y} - \text{LOD}) \qquad (4.1)$$

$$\hat{\sigma} = \sqrt{s^2 + \lambda\,(\bar{y} - \text{LOD})^2}$$

Where $\bar{y}$ and s^2 are the mean and variance of the detected values and λ is a correction factor tabulated by Cohen (1961), which is defined in function of the proportion of non-detect values and the gap between $\bar{y}$ and LOD value. This approach gave the following distribution for B(a)P levels of in raw river water (µg/L): N(4.2 10^{-3}, 2.3 10^{-2}). This distribution was used for reconstructing a long-term time series of contamination levels in raw river water over the period from 1993 to 2008.

Long- term time series of B(a)P levels in the air flow entering Paris region were given referring to the range of values (maximum: 4.0^{-6} mg/m^3, minimum: 1.0^{-6} mg/m^3) that is indicated in Quéguiner et al. (2010). Temporal variation of B(a)P concentration in the air flow were made up in such a way that the maximum and minimum values come in the middle of winter time and in the middle of summer time, respectively. The trend in which predominant air pollution of PAHs occurs in the wintertime is reported in Gusev et al. (2006). A sinusoidal function was then considered for long-term time series of B(a)P levels in air, given those high and low values:

$$C_{\text{air\_flow}}(t) = \frac{C_{max} + C_{min}}{2} + \frac{C_{max} - C_{min}}{2} \times \sin\left(2\pi\frac{t}{365} - \frac{\pi}{2}\right) \qquad (4.2)$$

where $C_{\text{air\_flow}}$ (t) (mg/m^3) is the time-dependent B(a)P concentration in air flow entering the target region, and C_{max} and C_{min} are the maximum and minimum B(a)P concentrations, respectively.

2-FUN environmental multimedia and PBPK models

The complete description of the 2-FUN multimedia environmental model consisting of freshwater, air, soil/groundwater, plants, and animal compartments is presented in the project deliverables (available at www.2-fun.org). Mass balance equations for each compartment are shown in the former section 2.2.

A PBPK model for men and women has been developed to simulate the body burden of various xenobiotics throughout the entire human lifespan, integrating the evolution of the physiology and anatomy from childhood to advanced aged. The complete description of the PBPK model is present in Beaudouin et al. (2010). That model is based on a detailed description of the body

anatomy and includes a substantial number of tissue compartments to enable the analysis of toxicokinetics for diverse chemicals that induce multiple effects in different target tissues. The model structure is identical for men and women. This model proposes a detailed compartmentalization of the human body with 22 organs. Only two compartments integrate several entities: the urinary tract (including the bladder, the ureters and the urethra) and the sexual organs (the testes, the epididymes and the prostate for men, and the ovaries, the fallopian tubes and the uterus for women). The lungs were separated into two compartments to distinguish the pulmonary functions from the lungs anatomy (the tissues). The main processes were modeled as follows:

- Absorption: Adsorption of PAHs occurs through different routes, but primarily occurs via food or water intake and via inhalation. For the purpose of this study, the intake was modelled as a direct input into the venous blood compartment, since most of the B(a)P ingested is assumed to be absorbed.

- Distribution: B(a)P is distributed into the various compartments by the blood flow and the partitioning from blood to tissues. All tissue compartments are assumed to be well-mixed and blood flow-limited. Distribution in tissue is then described by:

$$\frac{dQ_i(t)}{dt} = F_i \times \left(C_{art}(t) - \frac{C_i(t)}{PC_i} \right) \qquad (4.3)$$

where Q_i is the amount of chemical in compartment i (mg), C_i the concentration (mg/L), C_{art} the concentration in arterial blood (mg/L), F_i the blood flow entering in compartment i (L/min), and PC_i the tissue i - blood partition coefficient.

- Metabolism: Metabolism of B(a)P was limited to lungs and liver. The Michaelis-Menten relationship was used to describe B(a)P metabolism. The rate of metabolism (RAM) is then given by:

$$RAM_i(t) = \frac{Vmax_i \times C_{ven,i}(t)}{Km_i + C_{ven,i}(t)} \qquad (4.4)$$

where V_{max} is the maximal velocity (mg/min), K_m the Michaelis constant (mg/L) and $C_{ven,i}$ the concentration of B(a)P leaving the compartment i (mg/L) . The quantity of B(a)P in the liver and the lungs is obtained by:

$$\frac{dQ_i(t)}{dt} = F_i \times \left(C_{art}(t) - \frac{C_i(t)}{PC_i} \right) - RAM_i(t) \qquad (4.5)$$

- Excretion: Since B(a)P is highly metabolized, no excretion was assumed.

Integration of models in the Ecolego® platform

The multimedia and PBPK models were coupled on the common platform called Ecolego® (the information is available at www.facilia.se) to facilitate integrated full-chain assessment and to conduct the sensitivity and uncertainty analyses. An overall diagram depicting these coupled models is shown in Figure 4.1. Advanced methods for probabilistic and sensitivity analyses are available in the Ecolego system, such as: (i) Monte Carlo method for the propagation of parametric uncertainty; and (ii) several regression and Fourier approaches for conducting sensitivity analysis.

Parameterization for 2-FUN tool

Many input parameters are included in the 2FUN tool, and thus the parameterization for them is an important process to implement the tool. Some of these input parameters assumed to be highly uncertain. Giving probabilistic forms to those input parameters permits uncertainty and sensitivity analyses targeting investigated model outputs. The present study used probabilistic density function (PDF) to represent the probabilistic forms of those input parameters. As for the parameters defined in the environmental multimedia model, site-specific PDFs were obtained when actual monitoring datasets in a specific site (in this study, at the Seine River) are available. When there were no site-specific data for the parameters of interest, instead, global (generic) PDFs were derived based on the global datasets of the parameters.

The PDFs were assigned to several parameters that are assumed to be highly uncertain. These parameters were selected from those included in the compartments which are directly linked with the PBPK model through digestion of drinking water, foods, and milk (i.e., the freshwater, plant, and animal compartments). The detailed description about the probabilistic parameterization for plant and animal compartments is shown in a deliverable in 2-FUN project. It should be noted that the parameters contained in the air and soil/groundwater compartments were set by point estimates based on the information in relevant documents and expert judgments. All the main parameters used in the 2-FUN environmental multimedia model are listed with their PDFs or point estimates in Appendix of this thesis.

To simulate B(a)P internal dosimetry in humans throughout their lifetime, quantitative relationships were derived for the parameterization of the PBPK model to describe the changes of the human anatomy and physiology over their lifetime. The full description of the model parameterization is available in Beaudouin et al. (2010). Relationships were defined according to the data available (International Commission on Radiological Protection, 2002; Altman and Dittmer 1962; Lexell et al. 1988; Haddad et al. 2001; National Health and Nutrition Examination Survey 1995), and the associated inter-individual variability in key parameters was also assessed. For volumes, we modeled the changes during childhood and adolescence but also

the effects of aging on muscle (atrophy) and adipose tissues (increase in adulthood and decrease in advanced age). Due to the lack of data on blood flow changes with age, regional blood flows were assumed to change proportionally to tissue volumes, as proposed by Clewell et al. (2004). The adult fractional tissue blood flows referenced by ICRP (International Commission on Radiological Protection 2002) were used along with the age-specific tissue volumes and the adult tissue volumes. To maintain mass balance for the blood flows, the age-specific fractional blood flows were normalized, so they sum to one.

Metabolism was assumed to be dependent on age but not gender. The cytochrome P450 activity/expression in infants was related to the adult level given the functions developed by Johnson et al. (2006). The maximal velocity was then apportioned to the relative activity of the cytochrome P450 1A2 that displays the highest affinity towards B(a)P. The maximal velocities for adults were set to 4.4 mg/min in liver and to 0.00173 mg/min in lungs (Chiang and Liao 2006). The Michaelis constants were set to 1.36 mg/l in liver and to 0.06 mg/l in lungs (Chiang and Liao 2006). The blood-air partition coefficient was set to 590 (Chiang and Liao 2006). The tissue-blood partition coefficients were calculated based on published data (Poulin and Krishnan 1996; Poulin and Theil 2002).

Determination of influential parameters to model outputs

To perform a global sensitivity analysis and an uncertainty analysis for the defined model outputs in this case study, only the key parameters potentially influential to the model outputs were in advance selected and assigned their corresponding probability density functions (PDFs). The key parameters in each different compartment were independently chosen in the following ways:

- Parameter selection in freshwater compartment: A global sensitivity analysis was conducted for an intermediate model output, i.e., the concentration of B(a)P in drinking water (dissolved water sourced from freshwater compartment) (mg/m^3).
- Parameter selection in plant compartments: A global sensitivity analysis was conducted for an intermediate model output, i.e., sum of the concentration of B(a)P in each plant (root, potato, grain, leaf, and fruit) (mg/kg_{fw}).
- Parameter selection in animal compartment: A global sensitivity analysis was conducted for an intermediate model output, i.e., sum of the concentrations of B(a)P in beef (beef cow) and milk (milk cow) (mg/kg).
- Parameter selection in human compartment (PBPK model): A global sensitivity analysis was conducted for B(a)P concentrations in liver and lungs (mg/L), and the total quantity of metabolites formed in the liver and lungs (mg).

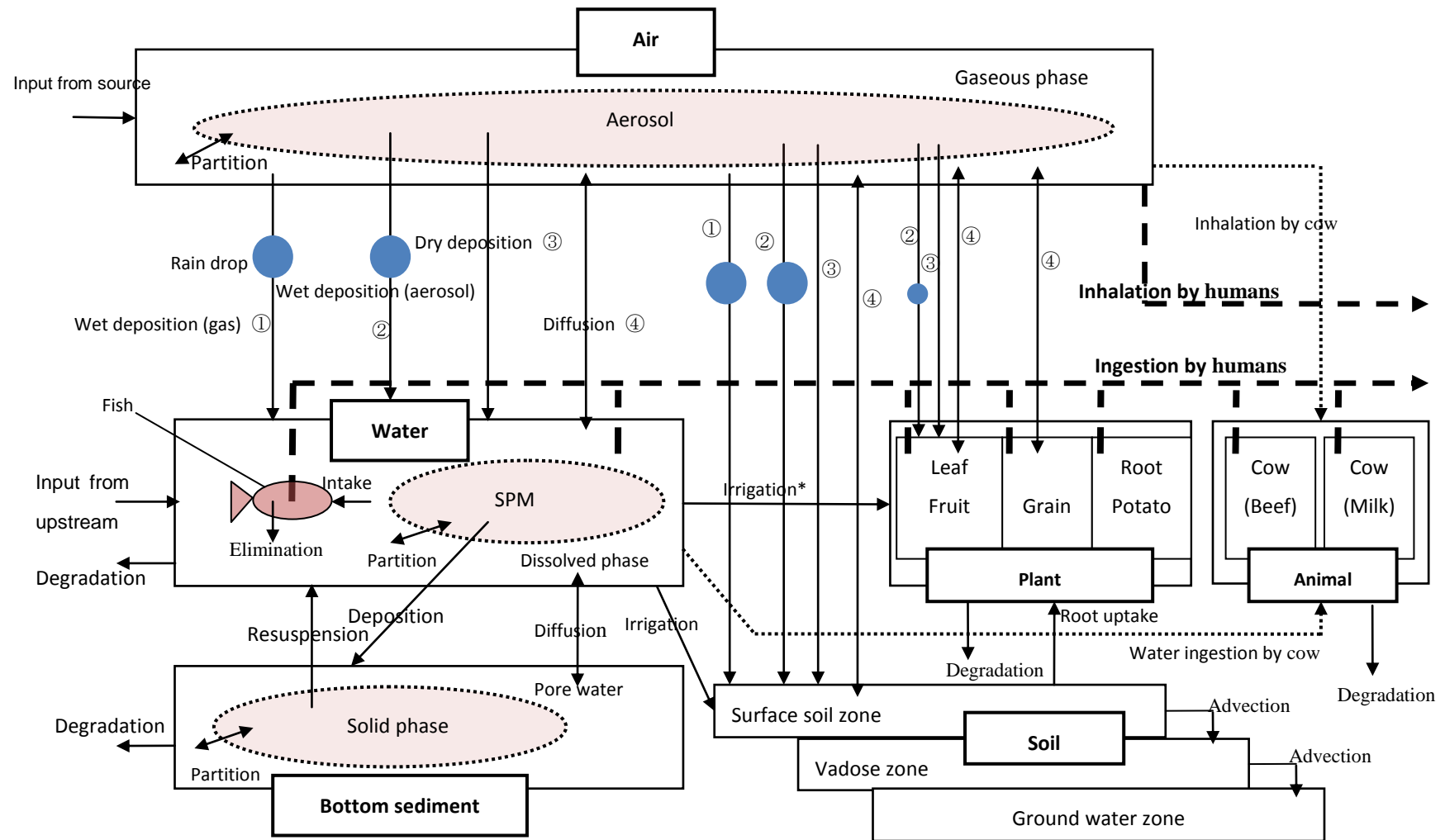All the selected parameters are presented in Table 4.1.

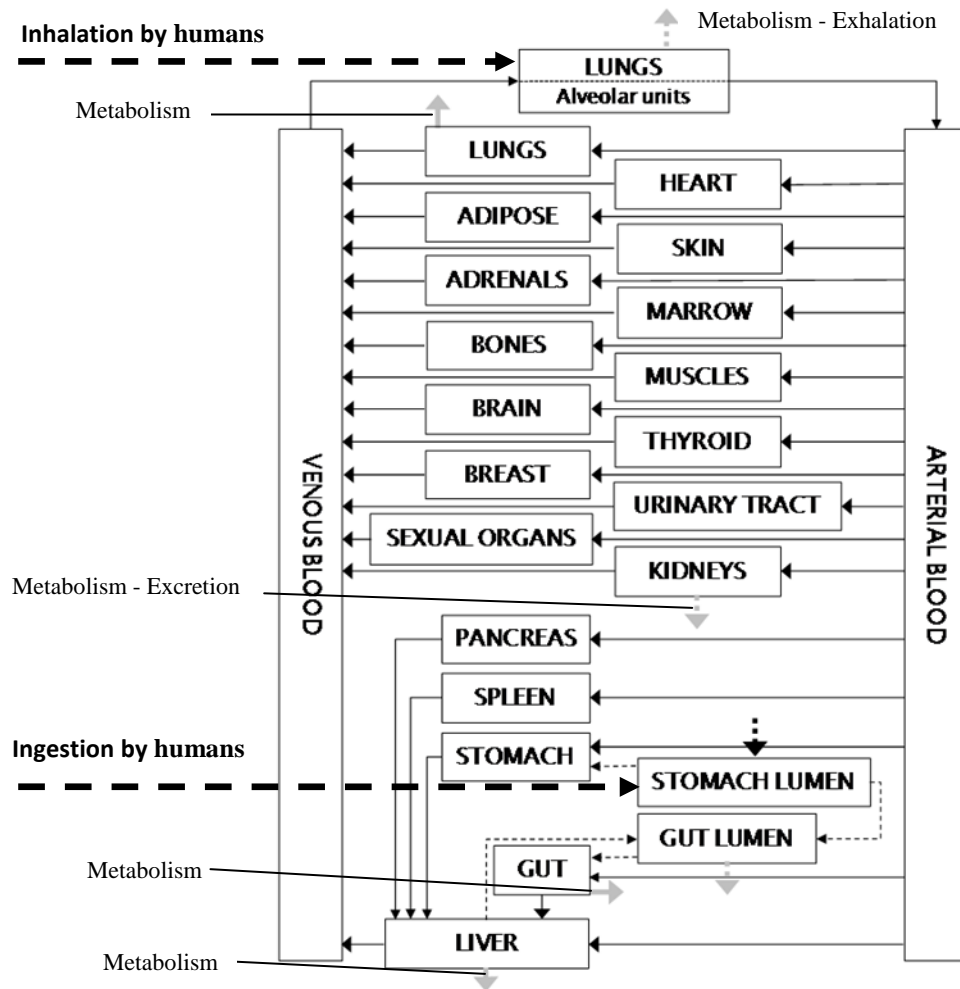Figure 4.1: Overall diagram of 2-FUN tool for the case study

Figure 4.1: (Continued)

Irrigation* indicates a chemical transfer to plants via irrigation water (only to leaf and grain compartments)

Table 4.1: The parameters used for the sensitivity and uncertainty analyses for the model outputs

Parameter	Unit	PDF	Relevant transport mechanism
Parameters in freshwater compartment			
1st empirical parameter for the rating curve relating suspended particulate matter (SPM) and flow rate in river (log(a))	-	N(-4.19, 0.33)	Concentration of time-dependent SPM (in freshwater compartment)
2nd empirical parameter for the rating curve relating SPM and flow rate in river (b)	-	N(0.99, 0.13)	Concentration of time-dependent SPM (in freshwater compartment)
Settling velocity of particles (W_c)	m/d	$LN_2(18.9, 3.0)$	Deposition from water to bottom sediments phase (in freshwater compartment)
1st empirical parameter for the relationship between time-dependent organic fraction in SPM and time-dependent SPM ($\alpha_{yoc,\,SPM}$)	-	Tri(0.15, 0.96, 0.55)	Time-dependent organic fraction in SPM (in freshwater compartment)
Parameters in plant compartments			
Octanol-water partition coefficient in grain compartment ($K_{OW,graiin}$)	-	$LN(10^{6.38}, 0.70)$	Chemical partitions between leaf tissue and water, between root and water, between pore water in soil and bulk soil, and etc
Leaf area per field area in leaf compartment (A_{leaf})	m^2/m^2		Chemical diffusion between leaf and atmosphere
Partition coefficient between air and water in leaf compartment ($K_{AW,leaf}$)	-	$Tri((0.86, 3.7, 1.78)10^{-4})$	Chemical partitions between leaf and air, between root and water, between leaf tissue and water, pore water in soil and bulk soil, and etc
Mass of leaf per field area in leaf compartment (M_{leaf})	kg_{fw}/m^2	LN(1.0, 0.017)	Chemical diffusion between leaf and atmosphere (only upward diffusion), and chemical transfer into leaf by leaf interception

Growth rate of leaf in leaf compartment ($K_{g, leaf}$)	d^{-1}	LN(0.035, 0.006)	Relative reduction of chemical in leaf by the growth of leaf
Parameters in animal compartment			
Lipid content of milk cow per bodyweight of milk cow (L_{milk})	kg/kg	Tri(0.06,0.31,0.21)	Partition coefficient between milk cow and water in milk cow
Bodyweight of milk cow (M_{milk})	kg	N(447.0, 76.0)	Chemical outflux by lipid excretion, lactation, urination and exhalation
Metabolism rate of cow ($K_{M, cow}$)	d^{-1}	Tri(5.5, 11.0, 8.25)	Chemical degradation by metabolism in both milk and beef cows
Parameters in human compartment (PBPK model)			
Percentage of the chemical absorbed (Abs) in human compartment	-	Uni(0.8, 1)	
Michaelis constant for metabolism in liver (Km_{liver}) in human compartment	mg/L	N(1.39, 1.39)	Chemical metabolism in liver
Michaelis constant for metabolism in lungs (Km_{lung}) in human compartment	mg/L	N(0.06, 0.06)	Chemical metabolism in lungs
Maximal velocity for metabolism in liver in human compartment ($Vmax_{liver}$)	mg/min	N(4.42, 4.42)	Chemical metabolism in liver
Maximal velocity for metabolism in lungs in human compartment ($Vmax_{lung}$)	mg/min	N(0.0017, 0.0017)	Chemical metabolism in lungs
Partition coefficient between blood and air in human compartment ($K_{blood, air}$)	-	N(590, 295)	Inhalation and exhalation of chemical

*N, LN, LN$_2$, Tri, and Uni indicate normal (mean, standard deviation), lognormal (mean, standard deviation), lognormal (geometric mean, geometric standard deviation), triangular (minimum, maximum, mode), uniform (minimum, maximum) distributions, respectively.

4.3 *Results and Discussion*

As described previously, the main model inputs in this case study are the long-term time series of B(a)P levels (from 1993 to 2008) in the river water at upstream zone and in the air flow entering the target region. Figure 4.2 and 4.3 show these time series of B(a)P levels.
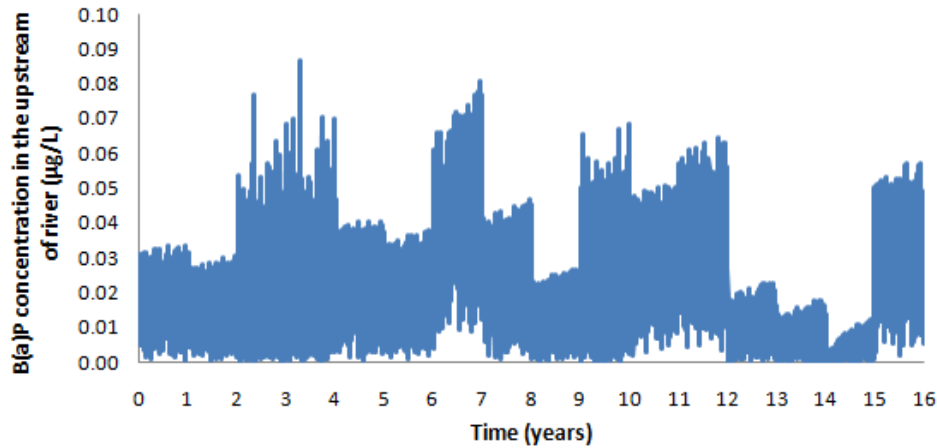

Figure 4.2: Time-dependent B(a)P concentration in the river water at upstream zone (μg/L) (Time 0 in horizontal axis corresponds to 1st January, 1993)

The deterministic and probabilistic calculations by the coupled models (2FUN-tool) were performed over the period corresponding to that for these model inputs. All the model outputs were given on monthly basis.
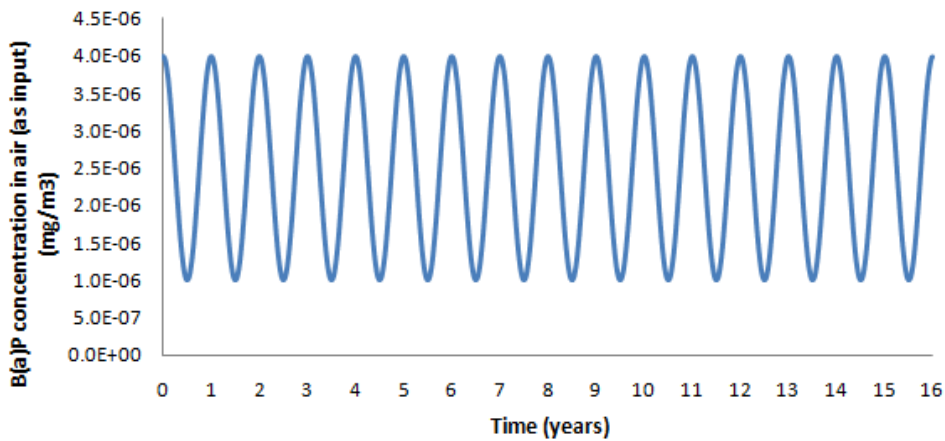

Figure 4.3: Time-dependent B(a)P concentration in the air flow entering the target region (mg/m^3)

The 2-FUN tool performed a probabilistic calculation for uncertainty/variability and sensitivity analyses. For this calculation, Monte Carlo approach randomly drew samples from the PDFs of parameters selected (Table 1) and generated the integrated distributions of specified outputs.

Quantity of the ingestion and inhalation of B(a)P through different pathways

Figure 4.4 presents the simulated quantities of ingested B(a)P via water, root, potato, grain, leaf, fruit, beef, and milk over the simulation period. This simulation result shows that the ingestion of leaf contributes the most to the exposure of B(a)P to humans, followed by fruit, milk ingestions. The quantity of B(a)P intake is calculated by multiplying the B(a)P concentration in each compartment by the corresponding daily ingestion rate (g/d). It was found in the result that the higher accumulation (concentration) of B(a)P in leaf is solely responsible for the higher quantity of leaf ingestion, considering that the daily ingestion rate of leaf is lower than those of other ingested foods and drinks except for fish. This simulation calculated the higher accumulation of B(a)P in leaf and fruit, probably because more pathways for B(a)P into leaf and fruit compartments are defined in 2-FUN tool than into other compartments (See Figure 4.1).

Figure 4.5 presents cumulative quantities of ingested (leaf, fruit, milk, and total ingestions) and inhaled B(a)P. The cumulative ingestion via leaf (0.20 mg) accounts for 66 % of cumulative total ingestion (0.30 mg) at the end of time, whereas the cumulative ingestion via fruit (0.067 mg) does 22 %. It indicates that about 90 % of total quantity of B(a)P ingestion is covered by the ingestions of leaf and fruit. The cumulative quantity of inhaled B(a)P (0.005 mg) accounts for only 1.7 % of total quantity of B(a)P ingestion at the end of time.
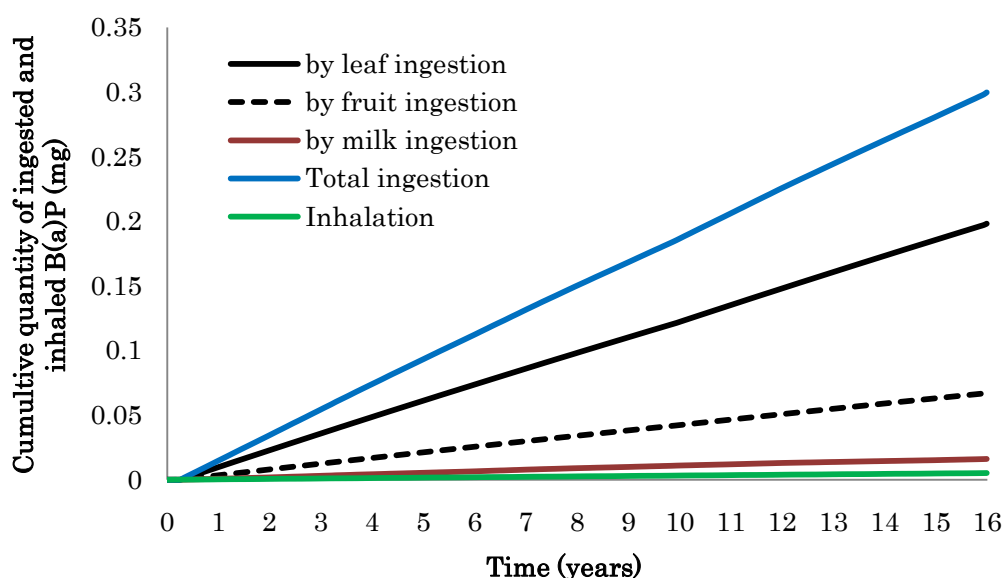

Table 4.5: Cumulative quantity of ingested and inhaled B(a)P (mg)
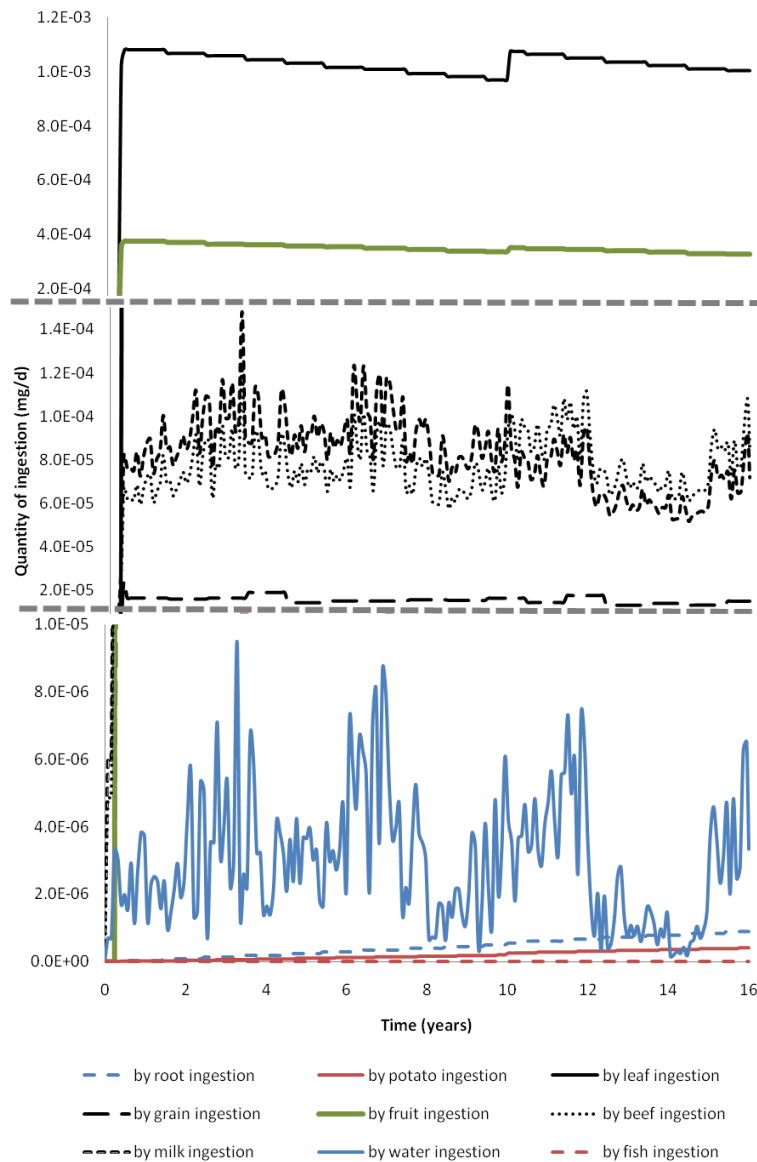
Figure 4.4: Time-dependent quantity of B(a)P entering humans through the ingestions of different food and drink items (mg/d)

Time series of B(a)P concentrations in human liver and lungs

Figure 4.4 presents B(a)P concentrations in human liver and lungs over the simulation period, with values at mean, 5th and 95th percentiles. It can be assumed that the values at 95th percentile represent 'pessimistic' scenarios in the context of health risk assessment. It was found that there is no clear accumulation of B(a)P in both liver and lungs over the period (after 2nd year). This can be explained by the fact that B(a)P is rapidly metabolized in those organs. The results show that the mean concentration in liver is almost five times higher than that in lungs over the

simulation period. Average ranges between B(a)P concentrations at 5th and 95th percentile over the simulation period, for liver, lungs, are 2 and 0.9 orders of magnitude, respectively. It indicates that the parametric uncertainties and variability contained in input parameter contribute significantly to propagation of such gaps in outputs.
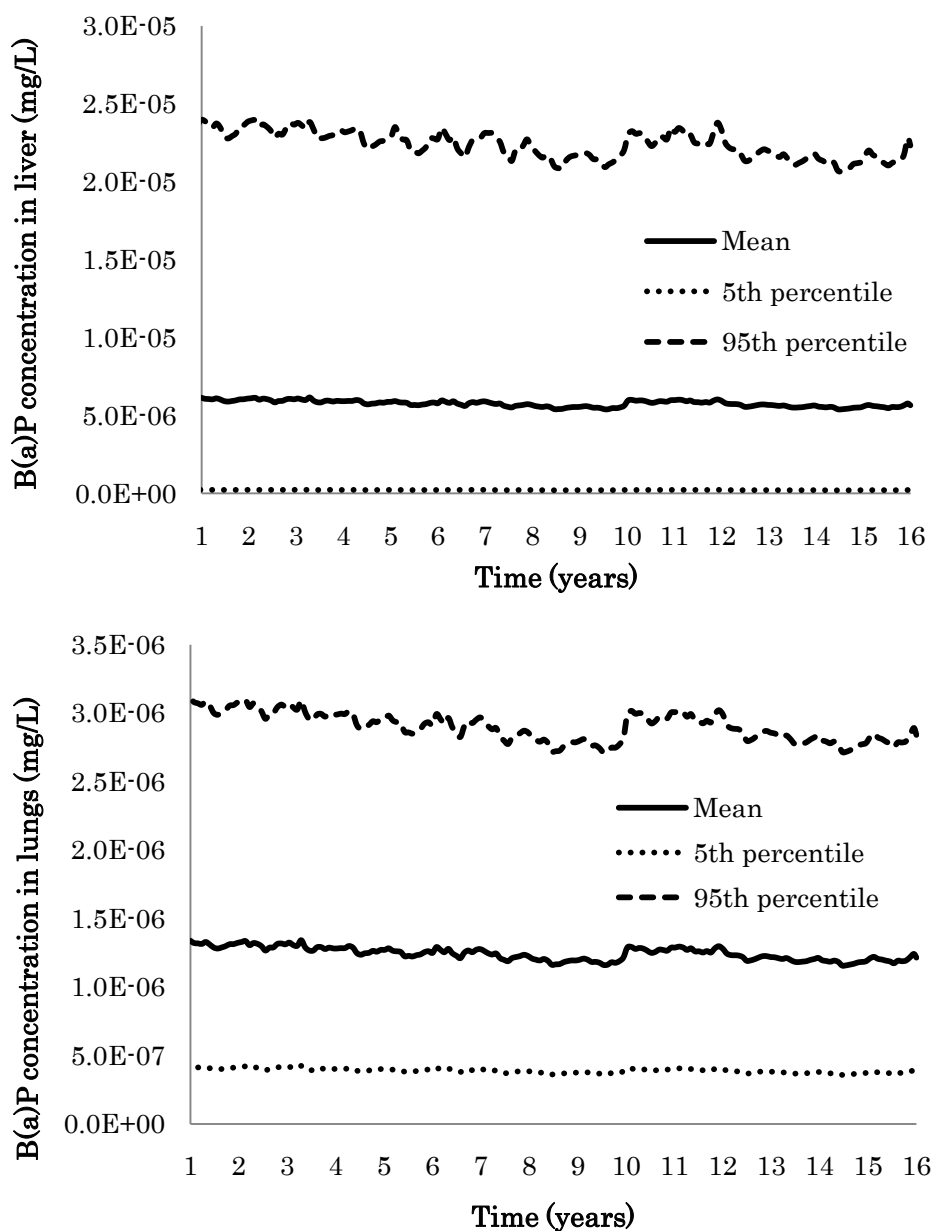

Figure 4.4: B(a)P concentrations in liver and lungs (mg/L) (Time 1 in horizontal axis corresponds to 1st January, 1994)

Cumulative quantity of metabolites formed in liver and lungs

Figure 4.5 shows the accumulated quantity of B(a)P metabolites formed in liver and lungs over the simulation period, with values at mean, 5th and 95th percentiles. The quantity of metabolites

can be an important indicator in health risk assessment as the toxicity of B(a)P is related to the formed metabolites. It can be said from the results that the accumulated quantity of metabolites formed in both organs is almost linearly increased over the simulation period. The ranges between the quantities of metabolites at 5th and 95th percentile at the end of time, for liver and lungs, are factors 2.0 and 80 (equivalent to 1.9 orders of magnitude), respectively.
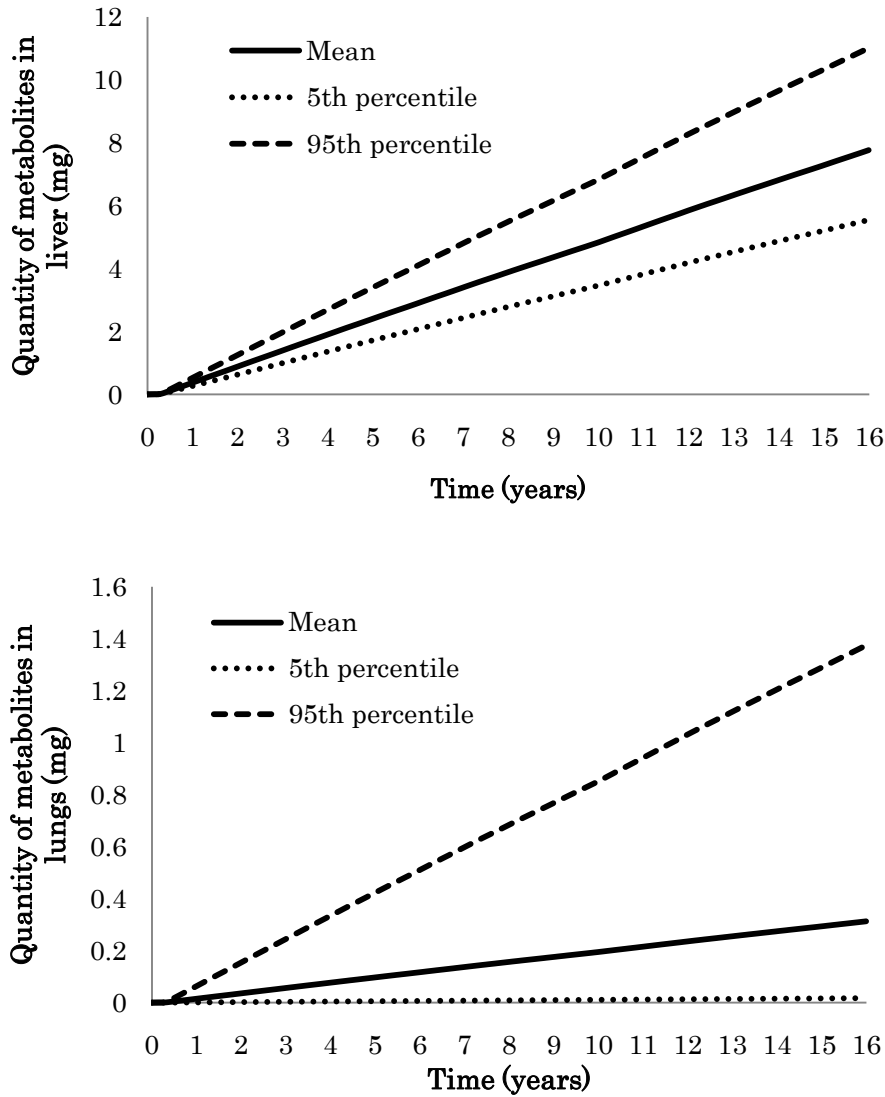

Figure 4.5: Quantity of metabolites formed in liver and lungs.

Global sensitivity analysis

A global sensitivity analysis was performed for specified model outputs in this study, i.e., the concentration of B(a)P in the liver and lungs, and the total quantity of metabolites formed in the liver and lungs (Figure 4.6) . The magnitude of sensitivity is shown by relative sensitivity index.

Figure 4.6 presents the average index value over the simulation period for each parameter. The result for each output is as follows:

- Concentration of B(a)P in liver: The most influential parameter is maximal velocity for metabolism in liver ($Vmax_{liver}$), followed by growth rate of leaf ($K_{g, leaf}$), and maximal velocity for metabolism in lungs ($Vmax_{lung}$).

- Concentration of B(a)P in lungs: Maximal velocity for metabolism in liver ($Vmax_{liver}$), followed by growth rate of leaf ($K_{g, leaf}$), and Michaelis constant for metabolism in lungs (Km_{lung}).

- Metabolites in liver: Partition coefficient between air and water in leaf ($K_{AW,leaf}$), followed by Octanol-water partition coefficient in grain ($K_{OW,graiin}$) and growth rate of leaf ($K_{g, leaf}$).

- Metabolites in lungs: Maximal velocity for metabolism in lungs ($Vmax_{lung}$), followed by Michaelis constant for metabolism in lungs (Km_{lung}) and liver (Km_{liver})
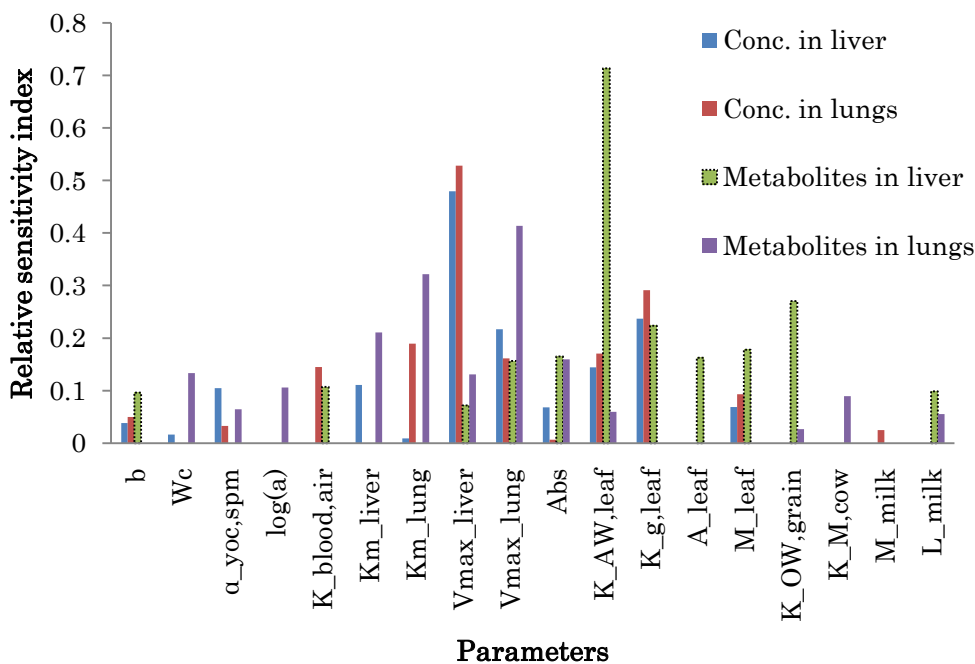
Figure 4.6: A global sensitivity analysis for specified model outputs

These results indicate that the parameters used in human and leaf compartments have relatively high sensitivity to all the specified model outputs. The high accumulation of B(a)P seen in leaf reflects the high sensitivity of leaf parameters to the model outputs.

The model outputs of 2-FUN tool are not explicitly able to indicate health risks (e.g., risks of cancers in liver and lungs) at this stage of model development. Nevertheless, it can be

concluded that 2-FUN tool has the potential applicability for health risk assessment, taking into account multiple exposure pathways via inhalation and ingestion.

4.4 *Conclusion*

This study demonstrated the feasibility of the integrated modelling approach to couple an environmental multimedia and a PBPK models, considering multi-exposure pathways, and thus the potential applicability of the 2-FUN tool for health risk assessment. The global sensitivity analysis performed in this study effectively discovered the input parameters and exposure pathways sensitive to the specified model outputs, i.e., B(a)P concentrations and accumulated quantities of metabolites in liver and lungs. This information allows us to focus on predominant input parameters and exposure pathways, and then to improve more efficiently the performance of the modelling tool for the risk assessment.

Nevertheless, the following issues should be considered in future studies to make the 2-FUN tool more practical and comprehensive for health risk assessments:

- To compare simulated results (internal B(a)P concentrations and accumulated quantities of metabolites in the target organs) with the correspoinding biomonitoring data, and to evaluate the model peformance
- To incorporate dose-response functions into 2-FUN tool to facilitate the evaluation of potential risks of human disease occurrances
- To add new compartmental media, e.g., the indoor compartment, which should be included to evaluate the exposure via inhalation of indoor air
- To add dermal intake as a main exposure pathway
- To add soil ingestion as a specific exposure pathway into children
- To conduct parameter estimations in depth (either for point estimates or probability density functions) for the parameters of interest included in air and soil/groundwater compartments, referring to a wide range of information source in local (e.g., Paris region) and global scale
- To consider processing procedures, such as smoking and drying, and cooking of food, which are commonly thought to be the major source of food contamination by PAH (EC 2002)

5 GENERAL CONCLUSION

The conceptual and theoretical aspects of 2-FUN tool are briefly summarized in Chapter 2. Presenting the general roles of mathematical models for health and environmental risk assessment of chemicals and the extensive review on existing modelling approaches, prominent features of the 2-FUN tool are well identified in Chapter 1.

Although it is still in process of development, the 2-FUN tool at this stage of development is capable to address overall scheme of environmental health risk assessment, except for the step of evaluating actual health risks by dose-response functions. Chapter 4 shows the feasibility of the integrated modelling approach to couple an environmental multimedia and a PBPK models, considering multi-exposure pathways, and thus the potential applicability of the 2-FUN tool for health risk assessment.

Oreskes et al. (1994) stated that numerical models can never be "verified", meaning as established as true, because they are necessarily imperfect representations of the real world, e.g., the complexity of environmental systems and human activities/customs. Nonetheless, model evaluation is a critical process of model development, through not only benchmarking the mathematical solution of a model or comparison with observations, but also sensitivity and uncertainty analyses. Since model inputs and parameters can never be known accurately, assessment of the effect of uncertainty and variability in these variables is inevitable if model outputs (predictions) are to be used for decision making by policy makers or risk assessors. Chapter 3 presents novel statistical approaches (i.e., Bayesian methods) to derive probabilistic density functions for the parameters used in the 2-FUN tool. Such probabilistic parameterization allows performing the uncertainty and sensitivity analyses.

The development of modelling approaches for environmental and health risk assessment and its relevant studies, e.g. parameter estimations, require a great deal of effort and interdisciplinary knowledge, and thus are time-consuming. It is essential that scientists and engineers from different scientific fields collaborate closely to establish more practical and comprehensive modelling approaches. Finally it can be concluded that the studies made for this dissertation would contribute to stimulating the study fields relevant for the environmental and health risk assessment and facilitating the progress of such study fields.

6 REFERENCE

Ahbe S, Braunschweig A, Müller-Wenk R (1990) Methodik für oekobilanzen auf der basis ökologischer optimierung. Schriftenreihe Umwelt. BUWAL Nr. 133, Bern.

Altman PL, Dittmer DS (1962) Growth, including reproduction and morphological development. Federation of American Societies for Experimental Biology, Washington, D. C.

Arain MB, Kazi TG, Jamali MK, Jalbani N, Afridi HI, Shah A (2008) Total dissolved and bioavailable elements in water and sediment samples and their accumulation in *Oreochromis mossambicus* of polluted Manchar Lake. *Chemosphere* 70: 1845-1856.

Asselman NEM (2000) Fitting and interpretation of sediment rating curves. *J Hydrol* 234: 228–248

Assmuth T, Hilden M (2008) The significance of information frameworks in integrated risk assessment and management. *Environmental Science & Policy* 11: 71-86.

Baccarelli A, Pfeiffer R, Consonni D, Pesatori AC, Bonzini M, Patterson Jr DG, Bertazzi PA, Landi MT (2005) Handling of dioxin measurement data in the presence of non-detectable values: Overview of available methods and their application in the Seveso chloracne study. *Chemosphere* 60: 898-906.

Balasubramanian S, Pappathi R, Raj SP (1995) Bioconcentration of zinc, lead and chromium in serially-connected sewage-fed fish ponds. *Bioresour Technol* 51: 193-197.

Ball DJ (2002) Environmental risk assessment and the intrusion of bias. *Environ Int* 28: 529-544.

Barber MC (2003) A review and comparison of models for predicting dynamic chemical bioconcentration in fish. *Environ Toxicol Chem* 22: 1963-1992.

Beaudouin R, Micallef S, Brochot C (2010) A stochastic whole-body physiologically based pharmacokinetic model to assess the impact of inter-individual variability on tissue dosimetry over the human lifespan. *Regul Toxicol Pharm* 57: 103-116.

Beaulieu SE, Sengco MR, Anderson DM (2005) Using clay to control harmful algal blooms: deposition and resuspension of clay/algal flocs. *Harmful Algae* 4: 123-138.

Blom G, Aalderink R (1998) Calibration of three resuspension/sedimentation models. *Wat Sci Tech* 37: 41–49.

Brandes LJ, Den Hollander H, Van de Meent D (1996) SimplBox 2.0: a nested multimedia fate model for evaluating the environmental fate of chemicals. RIVM Report no 719101029, RIVM, Bilthoven, The Netherlands.

Braunschweig A, Müller-Wenk R (1993) Ökobilanzen für Unternehmungen; eine Wegleitung für die Praxis. Verlag Paul Haupt, Bern.

Brock TCM, Alix A, Brown C, Capri E, Gottesbueren B, Heimbach F, Lythgo C, Schultz R, Streloke M (2009) Linking aquatic exposure and effects in the registration procedure of plant protection products (ELINK). *Proceedings,* 2nd SETAC Europe Special Science Symposium on Current developments on Environmental Risk Assessment for Plant Protection Products, Brussels, Belgium, September 17-18 2009: pp 55-66.

Burger J, Dixon C, Boring CS (2003) Effect of deep-frying fish on risk from mercury. *J Toxicol Env Heal A* 66: 817-828.

CEC (2003) A European environment and health strategy. Community strategy for environment and health (COM (2003) 338 final). Commission of the European Communities, Brussels, Belgium, June 2003. Available at http://europa.eu/legislation_summaries/environment/general_provisions/l28133_en.htm [Accessed 29 September 2010].

Chapman PM, Allen HE, Godtfredsen K, Z'Graggen MN (1996) Evaluation of bioaccumulation factors in regulating metals. *Environ Sci Technol* 30: 448-452.

Chiang KC, Liao CM (2006) Heavy incense burning in temples promotes exposure risk from airborne PMs and carcinogenic PAHs. *Sci Total Environ* 372: 64-75.

Ciffroy P, Decung F, Nouedoui L (2010b) Probabilistic fitting of sediment rating curves by a Bayesian approach. *In preparation for submission to a journal.*

Ciffroy P, Moulin C, Gailhard J (2000) A model simulating the transport of dissolved and particulate copper in the Seine river. *Ecol Model* 127: 99-117.

Ciffroy P, Siclet F, Damois C, Luck M (2006) A dynamic model for assessing radiological consequences of tritium routinely released in rivers. Application to the Loire River. *J Environ Radioactiv* 90:110-139.

Ciffroy P, Tanaka T, Marang L, Johansson E, Capri E (2010a) Simulation of time-dependent deposition/resuspension of contaminants to and from sediments in multimedia models: model and parametric uncertainty analysis. Submitted to *J. Soil Sed.*

Clarke B, Ghosh JK (1995) Posterior convergence given the mean. *Ann Stat* 23: 2116-2144.

Clewell HJ, Gentry PR, Covington TR, Sarangapani R, Teeguarden JG (2004) Evaluation of the potential impact of age- and gender-specific pharmacokinetic differences on tissue dosimetry. *Toxicol Sci* 79: 381-393.

Coetzee L, Du Preez HH, Van Vuren JHJ (2002) Metal concentrations in *Clarias gariepinus* and *Labeo umbratus* from the Olifants and Klein Olifants River, Mpumalanga, South Africa: Zinc, copper, manganese, lead, chromium, nickel, aluminium and iron. *Water SA* 28: 433- 448.

Cohen AC (1961) Tables for maximum likelihood estimates: singly truncated and singly censored samples. *Technometrics* 3: 535–541.

Cohen AC (1991) Truncated and Censored Samples: Theory and Applications, Marcel Dekker, New York.

Cullen AC, Frey HC (1999) Probabilistic techniques in exposure assessment. A handbook for dealing with variability and uncertainty in models and inputs. Plenum Publishing Corp., New York.

Curran KJ, Hill PS, Milligan TG, Mikkelsen OA, Law BA, Durrieu de Madron X, Bourrin F (2007) Settling velocity, effective density, and mass composition of suspended sediment in a coastal bottom boundary layer, Gulf of Lions, France. *Cont Shelf Res* 27: 1408-1421.

Davison AC, Hinkley DV (1997) Bootstrap methods and their application. Cambridge University Press, UK.

De Conto Cinier C, Petit-Ramel M, Faure R, Garin D, Bouvet Y (1999) Kinetics of cadmium accumulation and elimination in carp *Cyprinus carpio* tissues. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* 122: 345-352.

De Jesus Mendes PA, Maier I, Thomsen L (2007) Effect of physical variables on particle critical erosion shear stress: Hydrostatic pressure, slope and changes in water density. *Estuar Coast Shelf S* 75: 317-326.

DeForest DK, Brix KV, Adams WJ (2007) Assessing metal bioaccumulation in aquatic environment: The inverse relationship between bioaccumulation factors, trophic transfer factors and exposure concentration. *Aquat Toxicol* 84: 236-246.

Den Hollander HA, Van Eijkeren J, Van de Meent D (2004) SimpleBox 3.0. Report 422517 00x. RIVM, Bilthoven, The Netherlands.

Di Guardo A, Ferrari C, Infantino A (2006) Development of a dynamic aquatic model (DynA Model): estimating temporal emissions of DDT to Lake Maggiore (N. Italy). *Environ Sci Pollut Res Int* 13: 50-58.

Droppo IG, Walling DE, Ongley ED (2000) The influence of floc size, density and porosity on sediment and contaminant transport. In Stone M, eds, Role of erosion and sediment transport in nutrient and contaminant transfer, IAHS Publications No. 263, Waterloo, Ontario, Canada, pp 141-147.

Duboudin C, Ciffroy P, Magaud H (2004) Effects of data manipulation and statistical methods on Species Sensitivity Distributions. *Environ Toxicol Chem* 23: 489-499.

Dubus IG, Brown CD, Beulke S (2003) Sources of uncertainty in pesticide fate modeling. *Sci Total Environ* 317: 53-72.

Durrieu G, Ciffroy P, Garnier JM (2006) A weighted bootstrap method for the determination of Probability Density Functions of freshwater distribution coefficients (Kds) of Co, Cs, Sr and I radioisotopes. *Chemosphere* 65: 1308-1320.

d'Yvoire MB et al. (2007) Physiologically-based kinetic modelling (PBK modelling): Meeting the 3Rs agenda. The Report and Recommendation of ECVAM Workshop 63, ATLA 35: 661-671.

EC (2004) European Union System for the Evaluation of Substances 2.0 (EUSES 2.0). Prepared for the European Chemicals Bureau by the National Institute of Public Health and the Environment. RIVM Report no. 601900005, RIVM, Bilthoven, The Netherlands.

EC (2002) Opinion of the Scientific Committee on Food on new findings regarding the presence of acrylamide in food. SCF/CS/CNTM/CONT/4 Final, Brussel, Belgium.

EEA (2010) 10 messages for 2010 Agricultural ecosystems. Available at www.eea.europa.eu/publications/10-messages-for-2010-agricultural-ecosystems [Accessed 20 September 2010].

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, New York.

El Ganaoui O, Schaaff E, Boyer P, Amielh M, Anselmet F, Grenz C (2004) The deposition and erosion of cohesive sediments determined by a multi-class model. *Estuar Coast Shelf S* 60: 457-475.

Fettweis M (2008) Uncertainty of excess density and settling velocity of mud flocs derived from in situ measurements. *Estuar Coast Shelf S* 78: 426-436.

Fox JM, Hill PS, Milligan TG, Ogston AS, Boldrin A (2004) Floc fraction in the waters of the Po River prodelta. *Cont Shelf Res* 24: 1699-1715.

Gelfand AE, Hills SE, Racine-Poon A, Smith AFM (1990) Illustration of Bayesian inference in normal data models using Gibbs sampling. *J Am Stat Assoc* 85: 972-985.

Gelman A, Carlin JB., Stern HS, Donald BR (2004) Bayesian Data Analysis, Second edition. CRC Press, London, UK.

Gelman A, Hill J (2007) Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge, UK.

Ghosh JK (1988) The sorting hypothesis and new mathematical models for changes in size distribution of sand grains. *Indian J Geol* 60: 1-10.

Goedkoop M, Spriensma R (1999) The Eco-indicator 99. A damage oriented method for life cycle Impact assessment. PRé Consultants, Amersfoort.

Graham GW, Manning AJ (2007) Floc size and settling velocity within a Spartina anglica canopy. *Cont Shelf Res* 27: 1060-1079.

Gusev A, Ilyin I, Mantseva L, Shatalov V, Travnikov O (2006) Progress in further development of MSCE-HM and MSCE-POP models. Tech. rept. 4/2006. EMEP/MSC-E.

Gutièrrez S, Fernandez C, Barata C, Tarazona JV (2009) Forecasting risk along a river basin using a probabilistic and deterministic model for environmental risk assessment of effluents through ecotoxicological evaluation and GIS. *Sci Total Environ* 408: 294-303.

Haddad S, Restieri C, Krishnan K (2001) Characterization of age-related changes in body weight and organ weights from birth to adolescence in humans. *J Toxicol Env Heal A* 64: 453-464.

Hamilton SJ, Mehrle PM (1986) Metallothionein in fish: Review of its importance in assessing stress from metal contaminants. *Trans Am Fish Soc* 115: 569-609.

Hauschild MZ, Potting J (2005) Spatial differentiation in life cycle impact assessment – the EDIP2003 methodology. Guidelines from the Danish Environmental Protection Agency, Copenhagen.

Hauschild MZ, Wenzel H (1998) Environmental Assessment of products. Volume 2: Scientific Background. Chapman & Hall, London.

Heck T, Krewitt W, Malthan D, Mayerhofer P, Pattermann F, Trukenmuller A, Ungermann R (1997) EcoSense 2.0 - User's Manual. Universitat Stuttgart, IER.

Helsel DR (2006) Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it? *Chemosphere* 65: 2434-2439.

Helton JC (1993) Uncertainty and sensitivity analysis technique for use in performance assessment for radioactive waste disposal. *Reliab Eng Syst Safe* 42: 327-367.

Hertwich EG (1999) Toxic Equivalency: Addressing Human Health effects in life cycle Impact assessment. Thesis, University of California, Berkeley.

Hertwich EG, Mateles SF, Pease WS, McKone TE (2001) Human toxicity potentials for life cycle assessment and toxics release inventory risk screening. *Environ Toxicol Chem* 20: 928-939.

Hertwich EG, McKone TE, Pease WS (1999) Parameter uncertainty and variability in evaluative fate and exposure models. *Risk Anal* 19: 1193–1204.

Hill PS, Milligan TG, Geyer WR (2000) Controls on effective settling velocity of suspended sediment in the Eel River flood plume. *Cont Shelf Res* 20: 2095-2111.

Hofstetter P (1998) Perspective in life cycle Impact assessment. A structured approach to combine models of the technosphere, ecosphere and valuesphere. Ph.D. Thesis, Kluwer Academic Publishers, Dordrecht.

Huang YK, Lin KH, Chen HW, Chang CC, Liu CW, Yang MH, Hsueh YM (2003) Arsenic species contents at aquaculture farm and in farmed mouthbreeder (*Oreochromis mossambicus*) in blackfoot disease hyperendemic areas. *Food Chem Toxicol* 41: 1491-1500.

Huijbregts MAJ, Struijs J, Goedkoop M, Heijungs R, Jan Hendriks A, Van de Meent D (2005) Human population intake fractions and environmental fate factors of toxic pollutants in life cycle impact assessment. *Chemosphere* 61: 1495-1504.

International Commission on Radiological Protection (2002) Basic Anatomical and Physiological Data for Use in Radiological Protection: Reference Values. ElSEVIER.

IPCS (2001) Integrated risk assessment. Report prepared for the WHO/UNEP/ILO International Programme on Chemical Safety. WHO/IPCS/IRA/01/12. WHO, Geneva. Available at www.who.int/ipcs/publications/new_issues/ira/en [Accessed 29 September 2010].

IPCS, OECD (2003) Descriptions of selected key generic terms used in chemical hazard/risk assessment. International Programme on Chemical Safety Joint Project with OECD on Harmonisation of Hazard/Risk Assessment Terminology.

Ittekot V, Lane WP (1991) Fate of riverine particulate organic matter. In: Degens ET, Kempe S, Richey JE, Wiley J (eds), Biogeochemistry of Major World Rivers, SCOPE/UNEP Rep.42. New York: pp233-244.

Johnson TN, Rostami-Hodjegan A, Tucker GT (2006) Prediction of the clearance of eleven drugs and associated variability in neonates, infants and children. Clin Pharmacokinet 45: 931-956.

Jolliet O, Finnveden G, Frischknecht R, Giegrich J, Guinée JB, Hauschild M, Heijungs R, Hofstetter P, Jesen AA, Lindeijer E, Müller-Wenk R, Nichols Ph, Potting J, Wenzel C, White P (1996b) Impact assessment of human and eco-toxicity in Life Cycle Assessment; In "Towards a methodology for Life Cycle Impact Assessment", SETAC, Brussels.

Jolliet O, Margni M, Charles R, Humbert S, Payet J, Rebitzer G, Rosenbaum R (2003a) IMPACT2002+: A new life cycle impact assessment methodology. *J LCA* 8: 324-330.

Jolliet O, Pennington D, Amman C, Pelichet T, Margni M, Crettaz P (2003b) Comparative Assessment of the Toxic Impact of Metals on Humans within IMPACT2002. In: Dubreuil A (eds): Life Cycle Assessment of Metals – Issues and Research Directions, SETAC Press.

Khelifa A, Hill PS (2006) Models for effective density and settling velocity of flocs. *J Hydraul Res* 44: 390-401.

Kim Y, Sievering H, Boatman J, Wellman D, Pszenny A (1995) Aerosol size distribution and aerosol water content measurements during Atlantic Stratocumulus Transition Experiment/Marine aerosol and Gas Exchange. *J Geophys Res* 100: 23027-23038.

Kiss LB, Söderlund J, Niklasson GA, Granqvist CG (1999) New approach to the origin of lognormal size distributions of nanoparticles. *Nanotechnology* 10: 25-28.

Koning AD, Guinée J, Pennington D, Sleeswijk A, Hauschild M, Molander S, Nyström B, Pant R, Schowanek D (2002) Methods and typology report Part A: Inventory and classification of LCA characterisation methods for assessing toxic releases. OMNIITOX deliverable D11A.

Kozerski HP (2002) Determination of areal sedimentation rates in rivers by using plate sediment trap measurements and flow velocity – settling flux relationship. *Water Res* 36: 2983-2990.

Krewitt W, Mayerhofer P, Trukenmuller A, Friedrich R (1998) Application of the impact pathway analysis in the context of LCA. *J LCA* 3: 86-94.

Krewitt W, Trukenmueller A, Mayerhofer P, Friedrich R (1995) ECOSENSE: an integrated tool for environmental impact analysis. In: Kremers, H., Pillman, W. (Eds.), Space and Time inEnvironmental Information Systems. Umwelt-Informatik aktuell, Band 7. Metropolis-Verlag, Marburg.

Krone RB (1962) Flume studies of the transport of sediment in esturial shoaling processes. Final Report, Hydr Eng and San Eng Res Lab, University of California, Berkeley.

Krumbein WC (1938) Size Frequency Distributions of Sediments and the Normal Phi Curve. *J Sediment Res* 8: 84-90.

Kuttatharmmakul S, Massart DL, Coomans D, Smeyers-Verbeke J (2001) Comparison of methods for the estimation of statistical parameters of censored data. *Anal Chim Acta* 441: 215–229.

Lartiges BS, Deneux-Mustin S, Villemin G, Mustin C, Barres O, Chamerois M, Gerard B, Babut M (2001) Composition, structure and size distribution of suspended particulates from the Rhine River. *Wat Res* 35: 808-816.

Leppard GG, Flannigan DT, Mavrocordatos D, Marvin CH, Bryant DW, McCarry BE (1998) Binding of polycyclic aromatic hydrocarbons by size classes of particulate in Hamilton harbor water. *Environ Sci Technol* 32: 3633-3639.

Lexell J, Taylor CC, Sjostrom M (1988) What Is the Cause of the Aging Atrophy - Total Number, Size and Proportion of Different Fiber Types Studied in Whole Vastus Lateralis Muscle from 15-Year-Old to 83-Year-Old Men. J Neurol Sci 84: 275-294.

Lick W, Rapaka V (1996) A quantitative analysis of the dynamics of the sorption of hydrophobic organic chemicals to suspended sediments. *Environ Toxicol Chem* 15: 1038-1048.

Liu WC, Hsu MH, Kuo AY (2002) Modelling of hydrodynamics and cohesive sediment transport in Tanshui River estuarine system, Taiwan. *Mar Pollut Bull* 44: 1076-1088.

Luck M (2001) Projet radioécologie Loire : modélisation du transport des radionucléides sous forme dissoute et particulaire en Loire et Vienne. Report EDF/DRD/HP75/2001/049/A.

Mackay D (1991) Multimedia Environmental Models: The Fugacity Approach. Lewis Publishers, Chelsea, MI.

Mackay D, Joy M, Paterson S (1983) Quantitative water, air, sediment interaction (QWASI) fugacity model for describing the fate of chemicals in lakes. *Chemosphere* 12: 981-997.

Mackay D, Webster E, Cousins I, Cahill T, Foster K, Gouin T (2001) An introduction to multimedia models. Final report prepared as a background paper for OECD workshop Ottawa, October 2001.

Mantovanelli A, Ridd PV (2006) Devices to measure settling velocities of cohesive sediment aggregates: A review of the in situ technology. *J Sea Res* 56: 199-226.

McGeer JC, Brix KV, Skeaff JM, Deforest DK, Brigham SI, Adams WJ, Green A (2003) Inverse relationship between bioconcentration factor and exposure concentration for metals: Implications for hazard assessment of metals in the aquatic environment. *Environ Toxicol Chem* 22: 1017-1037.

McKone TE (1993a) CalTOX, a multimedia total exposure model for hazardous-waste sites, Part I: executive summary. UCRL-CR-111456PtI. prepared for the State of California, Department of toxic substances control. UCRL-CR-111456PtI. Lawrence Livermore National Laboratory, Livermore, CA, USA.

Mckone TE (1993b) CalTOX: a multimedia total-exposure model for hazardous-wastes sites: Part II: Multimedia transport and transformation model. prepared for the State of California, Department of toxic substances control. UCRL-CR-111456PtII. Lawrence Livermore National Laboratory, Livermore, CA, USA.

McKone TE, Hertwich EG (2001) The human toxicity potential and a strategy for evaluating model performance in life cycle assessment. *J LCA* 6: 106-109.

Mikkelsen OA, Hill PS, Milligan TG, Chant RJ (2005) In situ particle size distributions and volume concentrations from a LISST-100 laser particle sizer and a digital floc camera. *Cont Shelf Res* 25: 1959-1978.

Morehead MD, Syvitski JP, Hutton EWH, Peckham SD, (2003) Modeling the temporal variability in the flux of sediment from ungauged river basins. *Global Planet Change* 39: 95-110

National Center for Health Statistics (1995) Third National Health and Nutrition Examination Survey, 1988–1991. Selected Laboratory and Mobile Examination Center Data. Version 1, September, 1995, Hyattsville, MD, USA.

National Institute of Public Health and the Environment (1999) Environmental risk limits. RIVM601640001. Technical Report. Bilthoven, The Netherlands.

Oreskes N, Shrader-Frechette K, Belitz K (1994) Verification, validation and confirmation of numerical models in the earth sciences. *Science* 263: 641-646.

Partheniades E (1965) Erosion and deposition of cohesive soils. *J Hydraul Div proc ASCE* 91: 1-25.

Phillips DJH, Rainbow PS (1989) Strategies of trace metal sequestration in aquatic organisms. *Mar Environ Res* 28: 207-210.

Poulin P, Krishnan K (1996) A Tissue Composition-Based Algorithm for Predicting Tissue: Air Partition Coefficients of Organic Chemicals. *Toxicol Appl Pharmacol* 140: 521-522.

Poulin P, Theil FP (2002) Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. *J Pharm sci* 91: 129-156.

Power M, McCarty LS (1998) A comparative analysis of environmental risk assessment/risk management frameworks. *Environ sci technol* 32: 224-231.

Pufulete M, Battershill J, Boobis A, Fielder R (2004) Approaches to carcinogenic risk assessment for polycyclic aromatic hydrocarbons: a UK perspective. *Regul Toxicol Pharmacol* 40: 54-66.

Purkait B (2002) Patterns of grain-size distribution in some point bars of the Usri river, India. *J Sediment Res* 72: 367-375.

Quéguiner S, Genon LM, Roustan Y, Ciffroy P (2010) Contribution of atmospheric emissions to the contamination of leaf vegetables by persistent organic pollutants (POPs): Application to Southeastern France. *Atmos Environ* 44: 958-967.

Ramaswami A, Milford JB, Small MJ (2005) Integrated Environmental Modeling. Pollutant transport, Fate, and Risk in the Environment. John Wiley & Sons. Inc.

Rosenbaum R, Margni M, Jolliet O (2007) A flexible matrix algebra framework for the multimedia multipathway modeling of emission to impacts, *Environ Int* 33: 624–634.

Rosenbaum RK, Bachmann TM et al. (2008) USEtox-The UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in Life Cycle Impact Assessment. *J LCA* 13: 532-546.

Sanford LP, Maa J (2001) A unified erosion formulation for fine sediments. *Mar Geol* 179: 9-23.

Schüürmann G, Ebert RU, Nendza M, Dearden JC, Paschke A, Kühne R (2007) Predicting fate-related physicochemical properties. In Van Leeuwen CJ, Vermeire TG (eds) Risk Assessment of Chemicals: An Introduction, 2nd ed. Springer, Dordrecht, The Netherlands, pp 375-426.

Seuntjens P, Steurbaut W, Vangronsveld J (2006) Chain model for the impact analysis of contaminants in primary food products. Study report of the Belgian Science Policy.

Shah AQ, Kazi TG, Arain MB, Jamali MK, Afridi HI, Jalbani N, Baig JA, Kandhro GA (2009) Accumulation of arsenic in different fresh water fish species – potential contribution to high arsenic intakes. *Food Chem* 112: 520-524.

Sleeswijk AW, Heijungs R (2010) GLOBOX: A spatially differentiated global fate, intake and effect model for toxicity assessment in LCA. *Sci Total Environ* 408: 2817–2832.

Slotkin TA, Seidler FJ (2009) Benzo[*a*]pyrene impairs neurodifferentiation in PC12 cells. *Brain Res Bull* 80: 17-21.

Solomon K, Giesy J, Jones P (2000) Probabilistic risk assessment of agrochemicals in the environment. *Crop Prot* 19: 649-655.

Sorensen M, Autrup H, Moller P, Hertel O, Jensen SS, Vinzents P, Knudsen LE, Loft S (2003) Linking exposure to environmental pollutants with biological effects. *Mutat Res* 544: 255–271.

Stackelberg PE, Furlong ET, Meyer MT, Zaugg SD, Henderson AK, Reissman DB (2004) Persistence of pharmaceutical compounds and other organic wastewater contaminants in a conventional drinking-water-treatment plant. *Sci Total Environ* 329: 99-113.

Stackelberg PE, Gibs J, Furlong ET, Meyer MT, Zaugg SD, Lippincott RL (2007) Efficiency of conventional drinking-water-treatment processes in removal of pharmaceuticals and other organic compounds. *Sci Total Environ* 377: 255-272.

Sverdrup HV, Johnson MW, Fleming RH (1942) The Oceans. Prentice-Hall, New York.

Syvitski JP, Morehead MD, Bahr DB, Mulder T (2000) Estimating fluvial sediment transport: The rating parameters. *Water Resour res* 36: 2747-2760.

USEPA (1997) Guidance for cumulative risk assessment. Part1. Planning and Scoping. US Environmental Protection Agency, Washington, DC. Available at www.epa.gov/osa/spc/2cumrisk.htm [Accessed 29 September 2010].

USEPA (1997b) Waste Minimization Prioritization Tool (Beta Version 1.0): User's Guide and System Documentation. EPA530-R-97-019.

USEPA (2000) Guidance for Data Quality Assessment. EPA-QA/G-9.

USEPA (2002) Lessons learned on planning and scoping for environmental risk assessments. Planning Scoping Workgroup of Sci Pol Council Steering Committee. US Environmental Protection Agency, Washington, DC, January, 2002.

USEPA (2002) TRIM: Total Risk Integrated Methodology. Volume II: Description of Chemical Transport and Transformation Algorithms. EPA-453/R-02-011b. Washington, DC.

USEPA (2003) Framework for Cumulative Risk Assessment, Risk Assessment Forum. U.S. Environmental Protection Agency, Washington, DC, EPA/630/P-02/001F, May, 2003.

USEPA (2004) AQUATOX (Release 2). Modeling environmental fate and ecological effects in aquatic ecosystems. Volume 2: technical documentation. EPA-823-R-04-002.

USEPA (2007) Framework for Metals Risk Assessment. EPA 120/R-07/001. Washington, DC.

Van Zelm R, Huijbregts MAJ, Van de Meent D (2009) USES-LCA 2.0: a global nested multi-media fate, exposure and effects model. *J LCA* 14: 282-284.

Velcheva I (2002) Content and transfer of cadmium (Cd) in the organism of fresh-water fishes. *Acta Zool Bulg* 54: 109-114.

Vermeire T, Rikken M, Attias L, Boccardi P, Boeije G, Brooke D, Bruijn JD, Comber M, Dolan B, Fischer S, Heinemeyer G, Koch V, Lijzen J, Müller B, Murray-Smith R, Tadeo J (2005) European union system for the evaluation of substances: the second version. *Chemosphere* 59: 473-485.

Veyssy E, Etcheber H, Lin RG, Buat-Menard P, Maneux E (2004) Seasonal variation and origin of Particulate Organic Carbon in the lower Garonne River at La Reole (southwestern France). *Hydrobiologia*, 391: 113-126.

Vicente-Martorell JJ, Galindo-Riaño MD, Garcia-Vargas M, Granado-Castro MD (2009) Bioavailability of heavy metals monitoring water, sediments, and fish species from a polluted estuary. *J Hazard Mater* 162: 823-836.

Vinot I, Pihan JC (2005) Circulation of copper in the biotic comartments of a freshwater dammed reservoir. *Environ Pollut* 133: 169-182.

Voulgaris G, Meyers ST (2004) Temporal variability of hydrodynamics, sediment concentration and sediment settling velocity in a tidal creek. *Cont Shelf Res* 24: 1659-1683.

Wagner LE, Ding D (1994) Representing aggregate size distributions as modified lognormal distributions. *Transactions of the ASAE* 37: 815-821.

Warren C, Mackay D, Whelan M, Fox K (2007) Mass balance modelling of contaminants in river basins: Application of the flexible matrix approach. *Chemosphere* 68: 1232-1244.

Webster E, Mackay D, Di Guardo A, Kane D, Woodfine D (2004b) Regional Differences in Chemical Fate Model Outcome. *Chemosphere* 55: 1361-1376.

Wepener V, Van Vuren JHJ, Du Preez HH (2000) Application of the equilibrium partitioning method to derive copper and zinc quality criteria for water and sediment: A South African perspective. *Water SA* 26: 97-104.

WHO (World Health Organization) (2010) Health and Environment in Europe: Progress Assessment. ISBN 978 92 890 41680.

Winterwerp JC (1998) A simple model for turbulence induced flocculation of cohesive sediment. *J Hydraul Res* 36: 309-326.

Winterwerp JC (2007) On the sedimentation rate of cohesive sediment. *Proceedings in Marine Science* 8: 209-226.

Zhang Z, Lennox WC, Panu US (2004) Effect of percent non-detects on estimation bias in censored distributions. *J Hydrol* 297: 74-94.

Zhao C, Boriani E, Chana A, Roncaglioni A, Benfenati E (2008) A new hybrid system of QSAR models for predicting bioconcentration factors (BCF). *Chemosphere* 73: 1701-1707.

7 APPENDIX

Main input parameters in 2-FUN tool

1. Chemical specific parameters for Benzo(a)pyrene

Parameter (Unit)	2-FUN Symbol	Value (& PDF)	Source
Organic carbon–water partition coefficient (Logarithm) (L /kg)	$\text{Log}(K_{OC})$	6.01 (in log10)	1
Octanol-water partition coefficient (-)	K_{OW}	$10^{6.38}$ $\text{LN}(10^{6.38}, 0.70)$	2
Air water partition coefficient (-)	K_{AW}	$1.78 \ 10^{-4}$ $\text{Tri}((0.86, 3.7, 1.78) \ 10^{-4})$	3
Molar weight (g/mol)			4
Henry's law constant (P m^3 /mol)	H	0.046	1
Degradation rate in water (d^{-1})	$\lambda_{deg,water}$	0.3	4
Degradation rate in sediment (d^{-1})	$\lambda_{deg,sed}$	0.0006	4
Degradation rate in soil (d^{-1})	$\lambda_{deg,soil}$	0.003	4

1: Earl N, Cartwright CD, Horrocks SJ, Worboys M, Swift S, Kirton JA, Askan AU, Kelleher H, Nancarrow DJ (2003) Review of the Fate and Transport of Selected Contaminants in the Soil Environment. Draft Technical report P5-079/TR1, published by Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, UK.

2: Jager T, Rikken MGJ, van der Poel P (1997) Uncertainty analysis of EUSES: Improving risk management by probabilistic assessment. 6791020 039. RIVM Report. (http://rivm.openrepository.com/rivm/bitstream/10029/10274/1/679102039.pdf)

3: Hilal SH, Ayyampalayam SN, Carreira LA (2008) Air-liquid partition coefficient for a diverse set of organic compounds: Henry's law constant in water and hexadecane. *Environ Sci Technol*, 42: 9231-9236.

4: Mckone TE, Currie RC, Chiao FF, Hsieh PPH (1995) Intermedia Transfer Factors for Contaminants Found at Hazardous Waste Sites. Executive Report, a report prepared by the Risk Science Program, University of California, Davis for the State of California. Department of Toxic Substances Control, Sacramento, CA, US.

2. *Parameters in freshwater compartement*

Parameter (Unit)	2-FUN Symbol	Value (& PDF)	Source
1st empirical parameter for the relationship between time-dependent organic fraction in SPM and time-dependent SPM (-)	$\alpha_{yoc, SPM}$	0.55 Tri(0.15, 0.96, 0.55)	5
1st parameter of the linear relationship betweeen log(BAF) and log(K_{OW}) (-)	$\alpha_{0,fish}$	0.85	6
2nd parameter of the linear relationship betweeen log(BAF) and log(K_{OW}) (-)	$\alpha_{1,fish}$	-0.70	6
2nd empirical parameter for the rating curve relating SPM and flow rate in river (m/d)	b	0.99 N(0.99, 0.13)	5
Diffusion coefficient of oxygen in water	$B_{O2,water}$	0.000165	5
Biomass of fish in the river (g_{ww}/m^2)	$Biomass_{fish}$	20	E.J.
Boundary layer thickness above sediment (m)	δ_{sed}	0.015 Uni(0.0, 0.03)	5
Boundary layer thickness below sediment (m)	δ_{water}	0.00045 Uni((2.0, 7.0)10^{-4})	5
Maximum erosion rate ($g/m^2/d$)	e_{max}	2673 LN_2(2673, 5.4)	5
Daily ingestion rate (1) of fish by humans (kg_{fw}/d) (Age: 3 - 12)	$fish_{ing,rate}$	0.02	7
Daily ingestion rate (2) of fish by humans (kg_{fw}/d) (Age: 13 -)	$fish_{ing,rate}2$	0.021	7
Gravity acceleration (Pa m^2/kg)	g	9.8	
Depth of river (m)	H_{river}	3.0 LN(3.0, 1.0)	Default
Sediment consolidation rate constant (d^{-1})	λ_c	0.004	Sanford and Maa (2001)
Elimination rate constant in fish (d^{-1})	$\lambda_{elim,fish}$	0.001	E.J.
Length of river (m)	L_{river}	3000	Default
1st empirical parameter for the rating curve	Log(a)	-4.19	5

relating suspended particulate matter (SPM) and flow rate in river (-)		N(-4.19, 0.33)	
Molar mass of CO_2 (g/mol)	M_{CO2}	44	
Molar mass of water (g/mol)	M_{H2O}	18	
Molar mass of dioxygen (g/mol)	M_{O2}	32	
Manning's coefficient ($m^{1/3}$/s)	n_{Ma}	0.045 Uni(0.02, 0.07)	5
Power coefficient of the diffusivity relationship (-)	n_{water}	0.6 Uni(0.5, 0.7)	5
Porosity of sediment (-)	φ_{sed}	0.37 Uni(0.33, 0.41)	5
Universal gas constant ($Pa.m^3$/mol/K)	R	8.205	
Density of water (kg/m^3)	ρ_{water}	1000	
Critical shear stress for resuspension (N/m^2)	τ_r	0.18 LN_2(0.18, 3.78)	5
Settling velocity of particles (m/d)	W_c	18.9 LN_2(18.9, 3.0)	Tanaka et al. (2010)
Daily ingestion rate of tap water by humans (L/d^{-1})	$water_{ing,rate}$	0.283	8
Width of the river (m)	W_{river}	300	Default
Fraction of organic matter in sediment (-)	$y_{OC,sed}$	0.034 LN_2 (0.034, 2.25)	5

*N, LN, LN_2, Tri, and Uni indicate normal (mean, standard deviation), lognormal (mean, standard deviation), lognormal (geometric mean, geometric standard deviation), triangular (minimum, maximum, mode), uniform (minimum, maximum) distributions, respectively.

* E.J. denotes Expert Judgment

5: Ciffroy P, Tanaka T (2009) Intermediate database on parameters included in the freshwater sub-model. Prepared for 2-FUN project (Full-chain and Uncertainty Approaches for Assessing Health Risks in Future Environmental Scenarios), D2.3 (available at www.2-fun.org).

6: EC (1996) Technical Guidance Document in support of Commission Directive 93/67/EEC on risk assessment for new notified substances and Commission Regulation (EC) 1488/94 on risk assessment for existing substances. PARTS I, II, III and IV. Document. ISBN 92-827-8011-2. Luxembourg: Office for Official Publications of the European Communities.

7:   The data were obtained from the 2nd National Individual Survey on Food Consumption (INCA 2) conducted between the end of 2005 and April 2007, which contains food consumption data gathered over 7 days for over 4,000 participants, adults and children, living in mainland France (Available at www.afssa.fr)

8: Volatier JL (2000) Enquéte INCA Individuelle et Nationale sur les Consommations Alimentaires. Tec & Doc Lavoisier ed. Paris.

3. Parameters in air compartement

Parameter (Unit)	2-FUN Symbol	Value (& PDF)	Source
Height of atmosphere (m)	h_{atm}	10,000	E.J.
Rainfall scavenging ratio for gas (Ratio between the concentration in rainwater (mg/m3) and in gas (mg/m3)) (-)	$\Lambda_{scavenging,gas}$	45800	9
Rainfall scavenging ratio for particles (Ratio between the concentration in rainwater (in mg/m3 rainfall) and in aerosols (mg/m3 air)) (-).	$\Lambda_{scavenging,part}$	5440	9
Octanol-air partition coefficient for PAHs (-)	$Log(K_{OA,PAH})$	10.77	10
Total surface area of target region (m^2)	S_{atm}	1000,000	Default
Area of the region side (m^2)	$S_{frontier}$	1000,000	Default
Total suspended particles in atmosphere (g/m^3)	TSP_{atm}	$3.5 \ 10^{-5}$	11
Dry deposition velocity (m/d)	$V_{dry,atm}$	1680	12

9:   Sahu SK, Pandit GG, Sadasivan S (2004) Precipitation scavenging of polycyclic aromatic hydrocarbons in Mumbai, India. *Sci Total Environ* 318: 245-249.

10:   Finizio A, Mackay D, Bidleman T, Harner T (1997) Octanol-air partition coefficient as a predictor of partitioning of semi-volatile organic chemicals to aerosols. *Atmos Environ* 31: 2289-2296.

11.   Granier LK, Chevreuil M (1997) Behavior and spatial and temporal variations of polychlorinated biphenyls and lindane in the urban atmosphere of the Paris area, France. *Atmos Environ* 31: 3787-3802

12.   Wang XL, Tao S, Dawson RW, Wang XJ (2004) Uncertainty analysis of parameters for modeling the transfer and fate of benzo(a)pyrene in Tianjin wastewater irrigated areas. *Chemosphere* 55: 525-531.

4. Parameters in soil / groundwater compartement

Parameter (Unit)	2-FUN Symbol	Value (& PDF)	Source
Diffusion coefficient in surface soil and vadose zones (m^2/d)	D_{soil}	0.1	E.J.
Gas phase mass transfer coefficient at the interface between air and soil (m/d)	$D_{soil,atm}$	0.24	E.J.
Porosity in the aquifer (-)	ε	0.5	E.J.
Faction of organic carbon in soil (-)	f_{OC}	0.02	E.J.
Air content in soil (-)	G_{soil}	0.1	E.J.
Depth of root zone (m)	h_{root}	1.0	E.J.
Distribution coefficient in the aquifer (m^3/kg)	$K_{d,aquifer}$	1.0	E.J.
Bulk mass density of the aquifer (kg/m^3)	$\rho_{aquifer}$	2000	E.J.
Dry density of soil (kg/m^3)	$\rho_{soil,dry}$	1600	E.J.
Wet density of soil (kg/m^3)	$\rho_{soil,wet}$	2000	E.J.
Fraction of soil water storage at field capacity in the root zone (-)	θ_{fc}	0.3	E.J.
Fraction of soil water storage corresponding to the depletion fraction for no stress in plants (-)	θ_p	0.2	E.J.
Fraction of soil water storage at wilting point in the root zone (-)	θ_{wp}	0.15	E.J.
Maximum groundwater contribution (mm/d)	$Vu_{,max}$	1.0	E.J.
Water content in soil (-)		0.35	E.J.

5. Parameters in plant compartements

All the information about the parameters in plant compartments is present in the following document:

13: Legind CN, Trapp S (2010) Review of features, events and processes incorporated in existing models for chemical uptake into plants and animals – Proposal of the conceptual and mathematical 2-FUN model for assessing transfer of chemicals to plants and animals. Prepared for 2-FUN project, D2.4 (available at www.2-fun.org).

8 LIST OF ORIGINAL PUBLICATIONS

Original publications relevant for Chapter 3

[Papers on peer reviewed journals]

Tanaka T, Ciffroy P, Stenberg K, Capri E (2010) Regression approaches to derive generic and fish group-specific probability density functions of bioconcentration factors for metals. *Environ Toxicol Chem* 29: 2417-2425.

Tanaka T, Ciffroy P, Stenberg K, Capri E (2010) Probabilistic estimation of the settling velocity of particles using a Bayesian approach. Submitted to *Chemosphere*.

Ciffroy P, Tanaka T, Marang L, Johansson E, Capri E (2010a) Simulation of time-dependent deposition/resuspension of contaminants to and from sediments in multimedia models: model and parametric uncertainty analysis. Submitted to *J. Soil Sed.*

[Posters]

Tanaka T, Ciffroy P, Stenberg K, Capri E (2010) A statistical method to estimate the probabilistic uncertainty of the settling velocity of particles - a Bayesian method. SICA 2010 XXVIII Convegno Nazionale della Società Italiana di Chimica Agraria. September 20-21 2010, Piacenza, Italy.

Tanaka T, Ciffroy P, Stenberg K, Capri E (2009) Regression analyses to derive exposure-and species-specific probability density functions of bioconcentration factors for metals. SETAC Europe 19th Annual Meeting. May 31-June 4 2009, Gothenburg, Sweden.

Original publications relevant for Chapter 4

[Paper on peer reviewed journals]

Ciffroy P, Tanaka T, Johansson E, Brochot C (2010) Linking multimedia environmental and PBPK models to assess health risks of B(a)P associated to drinking water – A case study. Submitted to *Environ Geochem Hlth.*

[Poster and Abstract]
Brochot C, Tanaka T, Johansson E, Beaudouin R, Zeman F, Bois FY, Ciffroy P (2010) Integrating multimedia environmental models and effect models on a common platform for risk assessment. IUTOX 2010 XII International Congress of Toxicology. July 19-23 2010, Barrcelona, Spain [Awarded for '**ECETOC Young Scientist Award**' as best poster presentation recognition, given to Brochot C and Tanaka T].

Ciffroy P, <u>Tanaka T</u>, Maurau S, Brochot C, Johansson E, Roustan Y, Capri E 2010 Linking multimedia environmental and PBPK models on a common platform to assess health risks of PAHs : a case study including uncertainty/sensitivity analysis. SEGH 2010 International Conference and Workshops. June 27-July 2 2010, Galway, Ireland.