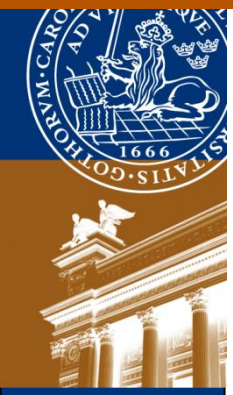

The impact of algal toxicity on life-cycle impact assessment of plastic additives and the potential of using QSAR predictions to fill the algae data gap

Julia Grönholdt Palm

2014

Environmental Science

Master thesis 30 ECTS

Lund University

The impact of algal toxicity on life-cycle impact assessment of plastic additives and the potential of using QSAR predictions to fill the algae data gap

Julia Grönholdt Palm

2014

Master thesis 30 ECTS, Lund University

Internal supervisor: Ullrika Sahlin, Centre for Environmental and Climate Research (CEC), Lund University

External supervisors: Tomas Rydberg and Magnus Rahmberg, IVL

Abstract

Purpose There is a need to find a quick way to assess the impacts of the growing amount of globally manufactured and emitted chemical substances. This paper evaluates the use of Quantitative Structure Activity Relationships (QSAR) for predicting environmental effects of plastic additives in Life Cycle Impact Assessment (LCIA). It also evaluates the impact on so called Characterization Factors (CF) when including toxicity on algae as opposed to only chordate and arthropod.

Method A review concluded that few (39) toxicity data for algae (experimental and QSAR predicted) were available for the 159 plastic additives of concern. To fill the data gap, a QSAR for algal toxicity was constructed that was able to predict toxicity for 54 substances. CFs were calculated and assessed based on; 1. QSAR predicted data for arthropod and chordate, 2. QSAR predicted data for arthropod, chordate and algae and 3. Experimental data for all three phyla.

Results and discussion CFs could be calculated considering algal toxicity for totally 97 out of the 159 substances. Algae were overall less sensitive to the substances leading to lower CFs when it was included. The correlation between the effect data of algae and the other two phyla was very small resulting in an altered internal rank when algal data was included.

Conclusions & recommendations

- The sensitivity of the species varied both between phyla and between substances.
- The inclusion of algal effect data did alter the internal rank of the resulting CFs although not extensively.
- Algae generally exhibited lower sensitivity to the additives. Not including algae in LCIA studies might therefore result in more conservative CFs.

Table of contents

Abbreviations

ACR	Acute/Chronic Ratios
AD	Applicability Domain
AiiDA	Aquatic Impact Indicator Database
BAF	Bioaccumulation factor
CDK	The Chemistry Development Kit
CF	Characterization factor (impact potential per kg emitted)
EC	European Commission
EC50	Effect Concentration (50% of test organism affected)
ECHA	European Chemical Agency
EPA	US Environmental Protection Agency
ETNCaq	Aquatic Exposure Threshold of no Concern
HC50	Hazardous Concentration for 50% of the included species
IS	Impact Score
LC50	Lethal concentration for 50% of test organisms
LCA	Life Cycle Assessment
LCI	Life Cycle Inventory
LCIA	Life Cycle Impact Assessment
MCS	Monte Carlo simulation
OECD	The Organisation for Economic Co-operation and Development
PAF	Potentially affected fraction of species
PLS	Partial least square
QSAR	Quantitative Structure-Activity Relationships
SMILES	Simplified Molecular Information Line Entry System
TEST	Toxicity Estimation Software Tool

1. Introduction

1.1 Background

Emissions of chemicals from anthropogenic sources have become a growing problem globally. To be able to regulate the usage of chemicals that causes the greatest threats there is a need to assess these emissions and their impact on humans and the environment. LCIA (lifecycle impact analysis) has become a commonly used tool where cumulative environmental impacts from all phases in a products lifecycle are assessed and characterization factors (CF) are derived representing the different types of impacts e.g. ecotoxicity. LCIAs make it possible to distinguish critical stages where the most significant emissions are from the initial raw material extraction and manufacturing process to the disposal, reuse or recycling (SAIC, 2006). However these models often demand an extensive amount of data of e.g. physiochemical properties and toxic effects, and empirical data is not always available. In these cases the use of data derived from QSAR-models (Quantitative Structure-Activity Relationship) for prediction of environmental fate and toxicity and could be an alternative to fill in the gaps where data is missing. With these models it is also easy to rapidly gather a vast amount of data which makes it a suitable tool to use for a first prioritization among a large set of substances. Usage and development of QSAR models also save monetary costs and reduces the need of animal testing (OECD, 2014). A lot of focus has been put on submitting data for e.g. environmental fate and human exposure while data for freshwater toxicity has received less attention and thus there are large gaps in the experimental ecotoxicological effect data, especially for new or less well-known substances (Henderson et al, 2011). Payet (2004) conducted a review of the data availability for calculation of Effect Factors for LCIAs where he assessed six of the largest databases for aquatic toxicity; Aquire, Pesticide Ecotoxicity Database (PED), IUCLID, Acute Toxicity Database (ATD), Fathead Minnow database (FMD), and ECETOC Aquatic Toxicity Database (EAT). Out of totally 113031 acute toxicity tests chordate represented 56% and arthropods 30%. Algal data represented only ca 5% with totally 5006 tests among all algal species. This highlights a great gap in the already scarce ecotoxicity data.

In the Swedish research program ChEmiTecs, emissions of additives from plastics in the Swedish societal material stock was assessed and the results projected that large quantities of additives are emitted each year through plastic migration (Westerdahl et al., 2010). It was estimated that about 2% are emitted to the environment every year which corresponds to approximately 50 000 tons. Some of these additives have been shown to have a negative impact on living cells e.g. by affecting the endocrine system (Stein, 2004). Despite this, in LCA literature additives are mostly disregarded and it is therefore hard to determine to what degree they contribute to the overall environmental impact. Van der Voet (2013) highlighted the importance of taking additives seriously and include them in LCI's as well as to improve LCIA databases by developing characterization factors for additives.

In a subsequent study, these additives were assessed using the LCIA tool 'USEtox' developed under the UNEP-SETAC Life Cycle Initiative. Due to the scarce amount of experimental toxicity data of plastic additives the applicability of using toxicity data derived from QSAR-models were evaluated (Rahmberg et al., 2012). This was done by comparing CFs derived using QSAR data to CFs derived using experimental data. However, only species from two phyla were included in the QSAR data due to lack of available QSAR models for algae and hence the obtained CFs were considered interim according to USEtox guidelines (Henderson, 2011). Therefore it was difficult to draw evident conclusions from the results.

In this study QSAR data of species from three different phyla will be used to meet the requirements in USEtox and EC's directive 67/548/EEC, Annex VI.

Species were chosen based on OECD guidelines; fish; 96-hour LC50, (OECD Test Guideline 203), a crustacea species; 48-hour EC50 (OECD Test Guideline 202) and an algal species; 72- or 96-hour EC50 (OECD Test Guideline 201). These are considered to best represent all aquatic organisms (OECD, 2013).

1.2 Purpose and Scope

The paper aims to assess the impacts of some of the plastic additives of concern in previous study by Westerdahl et al. (2010) while examining the compatibility of using QSAR predictions in USEtox. The purpose is also to investigate the availability of algal toxicity data, both QSAR and experimental, and to evaluate how substance characterization factors attained in USEtox is affected by including algal toxicity when calculating the ecotoxicological effect factor.

The intention is to answer the following questions:

1. How many of the substances of concern have available experimental toxicity data for algae, i.e. how large is the data gap?
2. Are there any QSAR models for algal toxicity that can be used to derive data for these substances and if not, is it possible to construct a QSAR model that is applicable to these substances?
3. For how many substances on the list can algal toxicity be predicted and how many substances have neither QSAR nor experimental values?
4. Which of the substances has the largest impact on freshwater ecosystems, based on CFs derived both with and without algal toxicity included?
5. How much and in what way does including data for algae affect the outcome when calculating CFs for freshwater toxicity in USEtox?
6. Do the differences in QSAR-predicted and experimental data result in large differences in assessed CFs and if so, are they more or less conservative?

The study comprises a list of 159 plastic additives, selected based on an earlier study where emissions of organic chemicals from consumer products containing plastic materials in Sweden were assessed (Westerdahl et al., 2010).

Characterization factors for freshwater aquatic ecotoxicological effects will be derived and the factor for impacts from emissions directly to freshwater will be used for the analyse. Since the aim of the paper concerns ecotoxicity the human toxicity impact are not included in this report.

Some of the substances on the list were identified as dissociating and due to this their physiochemical properties, environmental fate and ecotoxicity could not be predicted by the available estimation software programs. These substances were therefore not assessed further but were excluded from the dataset.

1.3 Environmental relevance

Ecosystems globally are exposed to chemical stress as result of large quantities of various substances being emitted everyday world wide. Many chemicals are not fully assessed before being commercially used which poses a threat to both humans and the environment. Plastic additives is an example of an organic substance group that has relatively recently received attention e.g. for being endocrine disruptive.

On the whole, this work will contribute to a better understanding of plastic additive's spread and effect in the environment and to clarify the extent to which different species are exposed and affected.

By developing a QSAR model and also compile data from existing models, it will facilitate and increase the efficiency of the evaluation and prioritization process which in turn will lead to a faster reduction and replacement of hazardous chemicals. Since the purpose of USEtox^{TM} is that it should be a free and globally accessible software the development of models and completion of data assists in the dissemination of information within and outside the EU. Production of QSAR models also helps to reduce the need for toxicity testing in animals.

2. Background

This section provides more detailed information of the underlying principles and components of the analytical methods (LCA, LCIA) and models (QSAR and USEtox) that will be employed and referred to in this paper. This part is meant for readers who are new to this area of studies and is intended to help in the understanding of the concept of LCA/LCIA and QSAR.

2.1 LCA and LCIA

Life-cycle assessments (LCAs) is a widely used tool to assess the environmental and human health impacts of products and processes and it has become the basis of EU's integrated product policy (IPP) (Hauschild, 2005). LCA uses a cradle-to-gate approach and as opposed to risk analysis strives to demonstrate relative differences between various options more than to quantify specific impacts in the system (EPA, 2006a). Therefore it is commonly used in comparative studies (Hauschild, 2005). It consists of four steps; goal definition and scoping, inventory analysis, impact assessment, and interpretation (EPA, 2006a).

In the impact assessment; LCIA (Life cycle impact assessment), the inventory data on input and output (resources, materials, emissions and waste) are evaluated based on environmental and human impact as well as resource consumption (Hauschild, 2005). The human and ecological health effects of each identified impact category are assessed and characterized (EPA, 2006a). The characterizations of chemical emissions are calculated by using multimedia models where several factors including environmental fate, exposure and effect are used. However, the CF from a model vary a lot depending on which model that is used. The models typically only cover CFs for around 1000 substances and therefore, when conducting a LCIA for a large amount of substances, the obtained CFs can vary a lot depending on which model has been used and some substances might lack CFs altogether (Henderson et al., 2011).

2.1.1 USEtox: Scientific consensus model

The variation in the CFs depending on which models are used contributes to the undermining of the reliability and comparability of the concept of characterization scores (Pant, 2004). To overcome this problem, in 2005, an extensive comparison of LCIA toxicity characterisation models was commenced by the United Nations Environment Program (UNEP)–Society for Environmental Toxicology and Chemistry (SETAC) Life Cycle Initiative.

The efforts resulted in a scientific consensus model, "USEtox", developed through harmonization between environmental fate models. It models the stages from emission to environmental fate, exposure and toxic effect in humans and the environment and derives characterization factors for

the assessed chemicals (Henderson et al., 2011). The CFs represents the impact per emitted unit of the substance of concern and combined with data of the mass emitted gives a total impact score (IS). This makes it possible to recognize the relative importance of separate emissions which is the main purpose of LCIAs (Mark Huijbregts and Tom McKone, 2010).
USEtox provides CFs for a large amount of substances and it is also possible to calculate CFs using the model matrix.

2.1.2 Limitations and uncertainty in LCIA modelling

There are several difficulties involved in LCIAs concerning choices of models, assumptions, absence of data and questionable quality of data.

Limitations that are often cited are; unclarity in what should be included (impact categories, stressors etc.), consistency within the categories when including severity/potency of stressors, only uncertainty related to data and the capacity of the model is assessed, difficulties in the interpretation of results and the weighting of impacts for compilation of results.

Compared to the more developed human risk-assessments in LCIAs there are more impact categories, more life-stages and a larger number of stressors which decreases the ability to produce results with high certainty, yet many practitioners do not perform uncertainty analyses. Because of the broad perspective of LCIAs, local impacts and risks are usually not assessed in detail. For chemicals this issue can typically imply that an LCIA may not account for background concentrations at specific locations but provide a more general picture of emissions during the life-cycle of a product (Bare, 2006). However an analysis by Hertwich et al. (1999) showed that compared to chemical, physical, and toxicity data parameters such as background concentrations represent an insignificant part of the uncertainty when assessing a large set of chemicals.

2.1.3 Uncertainties in the USEtox model

Rosenbaum et al. (2008) estimated uncertainty in USEtox based on an assessment between different LCIA models and found that the relative precision for the CFs are within a factor of 100-1000 for human health and 10-100 for freshwater ecotoxicity. This uncertainty range is just based on variation between the models not including uncertainty for the parameter input data.

Some uncertainty is related to the absence of valid mechanistic QSARs for estimation of substance properties. For freshwater ecotoxicity the sources of parameter uncertainty and variability are e.g. scarce data on bioconcentration factors for fish, chemical degradation rates and mostly; ecotoxicity effect data (e.g. extrapolations between chronic-acute data and assumption of linear dose-response curves). Further the application of homogenous compartments and the use of QSAR methods contribute to the uncertainty in USEtox (Jolliet and McKone, 2011).

There are freshwater toxicity CFs available for 2546 substances out of which 1247 are interim. These chemicals are flagged based on their properties where e.g. metals, dissociating or amphiphilic/surfactant substances are marked as interim. CFs are also classed as interim if the effect data does not cover at least three phyla (Rosenbaum et al., 2008). This is to make sure that the variations in physiology in different species are reflected and thereby ensure that the chemicals do not give a large variance in biological reactions (USEtoxTM, Hauschild & Larsen 2007). Rosenbaum et al. (2008) highlight that interim factors can be used but with precaution.

2.2 Quantitative structure–activity relationship

2.2.1 QSAR

QSAR models creates relations between chemical structures and physiochemical properties and these can be used to predict substances impact in the environment (Rodgers et al., 2011). The goal of QSAR modelling is to develop a mathematical expression that best describes this correlation between chemical properties and biologic responses (Eriksson et al., 2003). The QSAR model consists of a relationship between three elements; 1. a descriptor, 2. an endpoint that is being predicted (e.g. physiochemical property or biological activity), 3. the relationship between the two. Descriptors are structural properties that are used to describe molecular structures in the models (Walker, 2003). Common descriptors are e.g. molecular weight, K_{ow} and acid/base strength (pKa) etc. Training set data of substances with known descriptors and properties are first plotted in multiple regression models. These are then used to predict properties of query compounds with unknown properties (Rodgers et al., 2011).

2.2.2 Purpose and applications of QSAR

From the relationships between structure, chemistry and biology effect data achieved by QSAR methodology it is possible to construct predictable models that can be used within industry, academia or governmental agencies. Historically SARs and QSARs have typically been used by pharmaceutical companies and pesticide manufacturers to develop biologically active substances. In later years it has become an important tool for prioritizing among chemicals that have not yet been tested before making more costly experimental tests or to predict environmental fate, exposure and biological effect when experimental data is missing (Walker et al., 2002). It is also a useful instrument to predict combined effects of molecules which is useful to foresee what effects mixtures of chemicals could have in the environment or in formulations (Puzyn et al., 2010).

2.2.3 Principles for validation of QSARs

As a result of the complexity in methods to measure resemblance and accounting for multifaceted chemical behaviour the nature of QSAR modelling is complex. As the regulatory interest and acceptance for QSAR modelling grows it is important to make sure the construction and usage of the models are correct and that the validation process is transparent and objective. To help keep a solid scientific foundation in the development and use of QSAR in regulatory applications in 2004 the member countries of the OECD decided on five principles for validating QSAR models that are to be used for regulatory purpose (OECD, 2007). To assist inte the validation process a (Q)SAR model intended for regulatory purposes it should be coupled with the following information: 1. a defined endpoint; 2. an unambiguous algorithm; 3. a defined domain of applicability; 4. appropriate measures of goodness-of-fit, robustness and predictivity; 5. a mechanistic interpretation, if possible (OECD, 2004).

The OECD principles underline some of the main concerns in developing and using QSAR models which will be touched upon in the following section.

2.2.4 Developing QSAR models

In the original OECD document it is stated that “the intent of Principle 1 was to ensure clarity in the endpoint being predicted by a given model”. This includes providing thorough descriptions of test protocols used for deriving the measurements that were used in the training set data, including

factors affecting variability, knowledge and uncertainties (OECD, 2004). Also the quality of the endpoint data is essential since the data used in the training set regulates and defines the resulting QSAR (Walker, 2003). To produce a reliable model the measured data (y-variable) that are used to train the model therefore must originate from studies using equivalent methods, preferably from the same study (Rodgers et al., 2011). To make a QSAR for regulative purpose it is desired that the data used for the training set is derived using standardized test protocols such as OECD Test Guidelines, this will also speed up the validation process (OECD, 2007).

The second principle stresses that all algorithms that are used when producing a QSAR model are documented and submitted together with the QSAR model. This includes information on both the algorithm and the way the algorithm was developed. All stages in the modelling procedure should be specified so that the procedure is transparent and reproducible for the scientific community (OECD, 2007).

Third principle; "a defined domain of applicability" relates to the fact that QSAR models are limited to a certain group of chemicals for which it can make reliable predictions called the QSAR's applicability domain (AD). Netzeva et al. (2005) defined the AD as "The applicability domain of a (Q)SAR model is the response and chemical structure space in which the model makes predictions with a given reliability." The AD can be more or less constrained where a more constrained model often can make more accurate predictions for a smaller group of chemicals and vice versa. The information of the model AD helps the user to determine how reliable the prediction will be for a specific substance(OECD, 2007).

The purpose of the forth OECD principle is to assure that the performance of the models are tested. This part is also called the statistical validation and is made during the development of the model to find the optimal model complexity. The internal performance is assessed by looking at the goodness of fit and robustness of the model using only the training set data. The goodness of fit is determined by assessing how well the variance in the response training set is accounted for while the robustness is determined by looking at the stability of the parameters and hence stability of resulting predictions (OECD, 2007). The external performance of the model is tested by assessing predictions for new data that were not included in the training set (Eriksson et al., 2003, OECD, 2007).

The last principle on the OECD list urges the developer of the QSAR model to, if possible, include mechanistic interpretations to the model, meaning; if the model is consistent with other scientific knowledge in basic chemistry or toxicology this will be beneficial in the authentication process and shall therefore be documented. In the QSAR modelling mechanistic interpretation basically means to find relationships between descriptors and endpoints and to integrate a mechanistic and/or biological understanding of that relationship (OECD, 2007).

2.2.5 QSARs in LCIAs for screening and prioritization of chemicals

A LCIA serves as a screening-level impact evaluation for assessing the relative potential impacts of a system. Substances are characterized and the ones that warrant further investigations are usually evaluated in detail in qualitative risk screening analyses. However LCAs are often limited by the availability of quality data as well as by time constraints (collecting empirical data is time consuming). Since chemical toxicity data often are both scarce and involves relatively large uncertainties (e.g. fate calculations and extrapolations from animal-to-human and acute-to-chronic toxicity) chemical toxicity is often excluded from LCAs (Socolof, 2001).

When toxicity is included in LCIAs estimated data is typically used as an alternative to experimental data. Peer-reviewed data is preferred (e.g. HEAST, IRIS, HSDB), followed by other databases and literature and finally estimation methods like QSAR (Socolof, 2001, Pant, 2004). This order is also

applied in USEtox where experimental data is preferred, followed by specific QSARs (specific substance groups) and lastly generic QSARs (e.g. organic compounds). The uncertainty is generally decreased if this priority is applied (USEtox, 2013).

Over the last decade many countries have however adopted the QSAR methodology for chemical prioritization since it is a cost effective way to screen large inventories of chemicals.

3. Method

3.1 Data and implementation

This section describes the methods and tools that were employed in the study. It gives detailed information about how the data was gathered, and how the multimedia models were used to produce characterization-factors for the additives.

USEtox recommends that preference should be given to experimental data for physiochemical properties. However, in previous studies large variations were detected in experimental and estimated water solubility and it was reasoned that the datasets would be more comparable if only estimated data was used (Rahmberg et al., 2012). This reasoning was applied also in this study.

Toxicity data were gathered using several software tools and online databases described below. SMILES notations (Simplified Molecular Information Line Entry System) were used as input when searching the different databases to be sure that the same molecular structures were used.

3.1.1 Tools and databases

Physiochemical properties, degradation and bioaccumulation factor

- *EPI Suite* (Estimation Programs Interface) *4.11.*is a screening tool produced by EPA's Office of Pollution Prevention Toxics and Syracuse Research Corporation (SRC) and contains several physical/chemical property and environmental fate estimation programs
- *PBT Profiler* (Persistent, Bioaccumulative, and Toxic Profiles Estimated for Organic Chemicals) is an online screening tool developed by EPA for predicting chemicals tendency to persist in the environment, bio-concentrate in organisms, and be toxic (EPA, 2012)

QSAR predicted ecotoxicity data

- *TEST (Toxicity Estimation Software Tool) 4.1.* is based on the The Chemistry Development Kit (CDK) and estimates acute toxicity using QSAR methods.

Experimental ecotoxicity data

- *QSAR toolbox 3.2* (ECHA/OECD) integrates several databases for ecotoxicity including US-EPA Ecotox, Aquatic ECETOC, Danish EPA Database, Aquatic Japan MoE and ECHA CHEM (OECD, 2014a).
- *ECOTOX 4.0* is one of the largest databases and includes the previously independent databases AQUIRE, PHYTOTOX, and TERRETOX (EPA, 2014).
- *Aiida 3.0*. (Aquatic Impact Indicator Database) consists of the 70 largest worldwide aquatic ecotoxicity databases (Payet, 2013).

3.1.2 Physiochemical properties, degradation and bioaccumulation factor

Table 3.1 displays the required substance specific inputs parameters in USEtox. EPI Suite was used to derive the physiochemical properties and degradation parameters (air, soil water, sediment), and bioaccumulation factor in fish was attained from TEST 4.1 (consensus method).

For detailed description of how the parameters were derived see Appendix A.

Table 3.1 Required substance specific input parameters (Mark Huijbregts and Tom McKone, 2010)

Input parameter	Abbreviation	Unit	Column in substance data sheet	Necessary?[a]
Molecular weight	MW	$g.mol^{-1}$	4	Yes
Partitioning coefficient between octanol and water	K_{OW}	-	5	Yes
Partitioning coefficient between organic carbon and water	K_{oc}	$L.kg^{-1}$	6	No
Henry law coefficient (at 25˚C)	K_H25C	$Pa.m^3.mol^{-1}$	7	No
Vapour pressure (at 25˚C)	Pvap25	Pa	8	Yes
Solubility (at 25˚C)	Sol25	$mg.L^{-1}$	9	Yes
Partitioning coefficient between dissolved organic carbon and water	K_{DOC}	$L.kg^{-1}$	10	No
Degradation rate in air	$kdeg_A$	s^{-1}	16	Yes
Degradation rate in water	$kdeg_W$	s^{-1}	17	Yes
Degradation rate in sediment	$kdeg_{Sd}$	s^{-1}	18	Yes
Degradation rate in soil	$kdeg_{Sl}$	s^{-1}	19	Yes
Bioaccumulation factor in fish/biota	BAF_{fish}	l/kg	30	No

3.1.3 Ecotoxicity effect data

To be able to assess the impacts of including algae when calculating CFs and also to compare experimental data to QSAR data; three datasets of ecotoxicological effect data were compiled (table 3.2).

Table 3.2 Description of ecotoxicological datasets that were used to calculate CFs

Dataset	Content	Source
1	Effect data derived from QSAR models for two species from two different phyla; Chordate and arthropod	TEST.
2	Effect data derived from QSAR models for three species from three different phyla; Chordate, arthropod and algae	TEST, QSAR (model developed in present study)
3	Experimental effect data from three phyla	ECOTOX, QSAR toolbox, Aiida

Dataset 1

To derive predicted effect data TEST software tool was used. TEST was considered the best available QSAR-tool since it does not make predictions for substances out of a model's applicability domain and also have shown to give good predictions in comparative studies (Milan, 2012).

Effect data was predicted for common test species; fathead minnow (chordate) 96-hour LC50 (50% lethal concentration) and *D. Magna* (arthropod) 48-hour LC50 using the consensus method.

Dataset 2

Dataset 2 is composed of the same data as dataset 1 but QSAR-predicted data for algae species *Pseudokirchneriella subcapitata* (72-96 hour IGC50 (50% inhibitory growth concentration)) was

added as a third phyla. The algae data were derived from a QSAR model which was constructed for this study. The model, as well as the methods and tools that were used are described in section 4.2. *P. subcapitata* was used since it is known as a highly sensitive species and is therefore recommended as test species by OECD, EPA and ISO guidelines for legislative purposes when only one species is to be used (Aruoja, 2011). *P. subcapitata* also had the most complete set of measured data for the substances on the list.

Dataset 3

The third dataset was made out of experimental effect data for species from same three phyla as dataset 2. ECOTOX and QSAR toolbox were used for the first screening. From ECOTOX all data from the taxonomic groups "Fish", "Crustaceans" and "Algae, moss fungi" was extracted and from QSAR toolbox all aquatic taxonomic groups were included. The Aiida database was then used to fill in the gaps since it did not comprise a batch search function, e.g. where data for only one phyla was missing. Due to the short time limit of this study the original sources from which the experimental data originated were not examined in detail. In the short summaries provided by the Aiida database the data was rated depending on its reliability from 1-4 where 1 equals "reliable without restrictions". Sources marked 3 or lower were not included in this study. Data that were documented as, or did not appear reliable were excluded from the study.

3.2 Acute and chronic data

Most of the sources for experimental data gathered from both ECOTOX and QSAR toolbox did frequently not declare if the data were acute or chronic. To help determining what duration represent what type of test a table compiled by Payet (2004) was used (Table 3.3).

Payet (2004) reviewed the availability of toxicity data from some of the largest databases databases e.g. ECETOC, 2002; EU-Commission 2000; US-EPA 2001 and found that the availability was substantially higher for acute data than sub-chronic and chronic data.

Also in this study the major part of the experimental toxicity data that were found are acute. Likewise are the predicted data generated by the TEST tool acute (96h/48h LC50). Therefore, either would only chronic values extrapolated from acute data be used or these would be used where there were no chronic data available. Payet (2004) recommended in "Procedure for calculation of AMI Effect Factors " (s. 40), if the chronic data for a specific substance did not cover three phyla then chronic data would be derived using acute-to-chronic extrapolation factors. Since there were no substance for which three phyla were represented with chronic data all effect data was derived using an acute-to-chronic extrapolation factor. In USEtox the recommended extrapolation factor is set to 2 for organic substances based on a study by Rosenbaum et al. (2008).

Table 3.3 Time durations for different types of tests that were used to determine if retained data were respective acute, sub-chronic or chronic (Payet, 2004). ('*Based on guidelines from ISO, OECD, US-EPA. FIFRA, ASTM, UBA, and publications from Heger et Al, (Heger, Jung et al. 1995), ECETOC (ECETOC 1993), and the European Technical Guidance Document (EU-Commission 2002)).*

	Acute	Sub-chronic	Chronic [(1) (2)]
Vertebrates	Tests < 7 days	7 days Tests < 32 days	32 days Tests
Invertebrates	Tests < 7 days	7 days Tests < 21 days	21 days Tests
Plants	Tests < 7 days	-	7 days Tests
Algae	Tests < 3 days	-	3 days Tests

3.3 Ecotoxicological effect factor

In USEtox the ecotoxicological effect factor (EF) is determined by identifying HC50 (ΔPAF = 0.5) in a concentration- response relationship (figure 3.1). The HC50 is the average of the species-specific EC50 data and represent the concentration at which half of the species are exposed to concentrations higher than their EC50. The effect factor; EC50 is used due to its robustness which is considered to be critical in comparative studies such as LCAs instead of the more sensitive NOEC or HC5 (Rosenbaum et al., 2008). The EF is calculated using EQ 1.

$$\mathbf{EF} = \left(\frac{\mathbf{0.5}}{\mathbf{HC50}}\right) \qquad \text{(Eq. 1)}$$

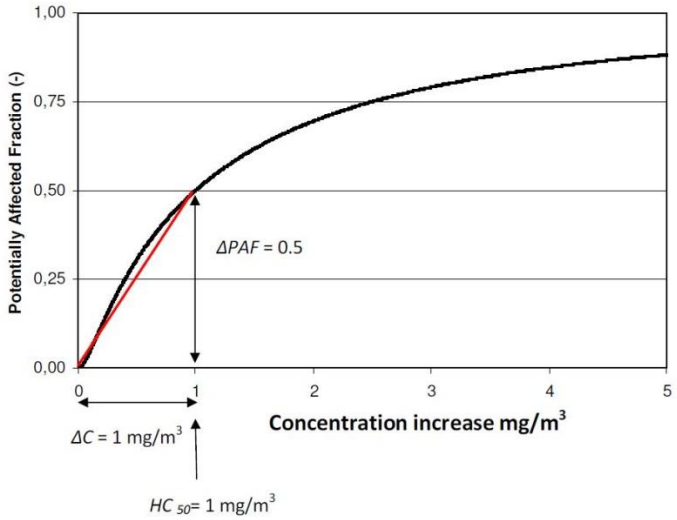

Figure 3.1 The figure displays how the EF factor is derived from the concentration-response curve (Huijbregts et al. 2010)

USEtox employs the AMI method developed by Payet (2004). In accordance with this method the species-specific EC50 is derived using the geometric mean of all effect data gathered for a specific species. This is a common method to use when deriving means of populations since it draws the extreme values towards the middle of the data, making it a more robust by making it less sensitive to outliers (Hauschild, 2007).

USEtox uses toxicity factor logHC50 to represent toxicity which is calculated using the equation provided in the USEtox manual (EQ 2.) (ns = number of species) (Mark Huijbregts and Tom McKone, 2010).

$$\mathbf{log\ HC50} = \frac{\mathbf{1}}{\mathbf{n_s}} * \sum_s \mathbf{logEC50_s} \qquad \text{(Eq. 2)}$$

In this study for the experimental effect data in dataset 3 the HC50 was derived by first calculating the geometric mean of the predicted EC50s (mg/l) for each species. Since only acute data were used both for experimental and predicted data the default acute-to-chronic extrapolation factor; 2, was used to derive the chronic-equivalent EC50 for all species (see section 3.1.5). The geometric meanEC50 were then logarithmized and the average of the logs were calculated to derive a logHC50 (mg/l) value for each substance. This value was implemented in the $av_{logEC50}$-colum in the USEtox template.

The same procedure was used for dataset 1 & 2 except there were only one species per phyla (48-hour *Daphnia magna*, 96-hour fathead minnow and for dataset 2 also; 72-96-hour *P. Subcapitata*) and thus HC50 was derived simply by calculating a mean of the three EC50s.

3.4 Uncertainty analysis

An uncertainty analysis was performed to evaluate how large change in CF one could expect by adding algae and to get a clue of what impact adding actual algae values may have on the assessment of CFs. Monte Carlo simulation (MCS) was performed in "R" v3.1.1. where data for *P. Subcapitata* were sampled from the experimental training set data that were used to construct the QSAR model. The samples were imported to the USEtox model which is implemented in Excel and then exportet back to R for the analyis. The sampling was made with 20 iterations and each substance were assigned a new sampled value. It is acknowledged that 20 iterations may be too small, but it was here considered sufficient to detect trends in this analysis.

4. Results

4.1 Review over algal toxicity data

Experimental algae toxicity data based on any algal species were found for totally 39 out of the 159 substances of concern. No existing QSAR model was found that were applicable for most substances of concern and it was decided to construct a QSAR to predict algal toxicity for these substances. The QSAR model that was trained on experimental data of algal species; *P. Subcapitata,* available from the substances on the list and from other substances. The QSAR was able to predict algal toxicity for 54 out of the 159 substances, leaving in total 66 substances with neither experimental nor QSAR toxicity data on algae.

4.1.1 Overview of available experimental data

The experimental data on algal toxicity were scarce on all databases that were used (QSARtoolbox, ECOTOX, Aiida). When EC50 (24-96h) for all algal species were included in the search; data for 16 substances were found in ECOTOX and 24 were found in QSARtoolbox (of which 16 were also found in ECOTOX). Data for totally 36 substances were found for arthropods and for chordate totally 43. The Aiida database was only used to fill gaps, yet Aiida included algae data for 13 substances that were not in either of the other databases and therefore appeared to be the most comprehensive database.

4.1.2 Overview of available QSAR models

QSAR models for predicting algal toxicity of the compounds were very scarce.

ECOSAR which is a software included in EPIsuite had models applicable for only 3 of the substances in the list; Bisphenol A (80-05-7), 1,2,3,6-Tetrahydro-N-(trichloromethylthio)phthalimide (133-06-2) and Butyl benzyl phthalate (85-68-7) belonging to the chemical groups poly phenols, thiophthalimides and esters.

TEST had no QSAR models for algal toxicity but included models for the uniform ciliate *Tetrahymena pyriformis. T. Pyriformis* is a unicel organism belonging to the group ciliates which in turn belongs to the protozoan phylum. Ciliates feed on bacteria and algae and also have no photosynthesis (Lynn, 2011). A study by (Schafer et al., 1994) demonstrated different sensitivity to toxins for algae (*Chlamydomonas reinhardi* and *Scenedesmus subspicatus*) *and T. Pyriformis*. For this reason, it was decided that the *T. Pyriformis* could not be used as a substitute for algae in dataset 2.

QSARtoolbox includes a number of external databases for QSAR models such as ECOSAR, Danish EPA database, Multicase Inc. etc. However there is no simple batch function that makes it possible to make predictions for a large number of chemicals. It is possible to make categorizations which can be saved as QSAR models and be applied on data within that domain. Yet, if the list includes chemicals from various chemical groups, as in this paper, this can be time demanding. It also requires some knowledge about chemical grouping.

4.2 Construction of QSAR model for algal toxicity

The model was trained using experimental data of algal toxicity for 35 chemicals from the list derived from QSAR toolbox, ECOTOX 4.0 and Aiida 3.0 as well as data for 45 related chemicals gathered in a study by (Furusjö et al., 2005). The measured endpoint was EC50 *P. Subcapitata* (72-96h).

1. To transform SMILES codes to sdf. Open Babel 2.3.2 was used
2. Discovery studio 4.0 "Clean Geometry" function was employed to optimize the geometry of the structures by accounting for element types, bond orders, number of bonds, and valences.
3. Hydrogen atoms were added to the structures using Open Babel 2.3.2.
4. To calculate molecular descriptors Dragon 6.0 software was used (Todeschini et al., (undated)). The descriptors were imported to SIMCA together with the ecotoxicity data of *P. Subcapitata.*
5. In SIMCA, Partial least squares regression (PLSR) was performed to examine the relationship between the x-values (descriptors) and the y-values (biological responses) as proposed by (Lindgren et al., 1996).

Table 4.1 Description of the methods that were used to develop and validate the model

Endpoint	EC50 *P. Subcapitata* (72-96h)
Algorithm	PLSR
Molecular descriptors	Dragon 6.0 (1D, 2D, 3D)
Applicability domain	DModX/PModX
Validation	Internal: R^2, Q^2, RMSEE

4.2.1 PLS

In this study PLS Regression (partial least squares) method was used to examine the relationship between the x-values (descriptors) and the y-values (biological responses) as proposed by (Lindgren et al., 1996). One important advantage of the PLS model is that it can treat datasets that contains more variables than observations which MLR (multi linear regression) cannot. It typically uses several hundreds to thousands of descriptors. Since the experimental data for algal toxicity was limited for the chosen substances, there were only a small set of observations on which the QSAR model could be based and therefore the PLS model was suitable in this study. Latent vector models like PLS also makes it easier to graphically display results which facilitate the interpreting of the results (Lindgren et al., 1996).

4.2.2 Applicability domain

The applicability domain was tested using SIMCAs DModXPS (Distance to Model in X space for the Prediction Set) and DModXPS+ that also take into consideration the distance in the model plane. From these measures the probability that a new (predicted) substance belongs to the model PModXPS+ was determined. Substances with a probability of less than 5% (PModXPS+ <0.05) was excluded.

4.2.3 Validation

For validation of the model three quality measures were used: goodness of fit, predictive ability and performance.

Goodness of fit

R^2 is a measure of fit and shows how well the model explains the training data. It expresses the fraction of the sum of squares that is explained by the model.

Predictive ability

The predictive ability was assessed using PLS cross validation (internal validation).
A 7th of the data was systematically excluded and a model was calibrated for the remaining data whereupon the model was validated based on the predictions for the excluded data.
Q^2 is the percent of the total variation of the Y that can be predicted by a component (Eq. 3).

Q^2= (1.0 - PRESS/SS) (Eq. 3)

Prediction error sum of squares (PRESS) is the (squared) difference between predictions and observed values for the excluded Y-data. For each component PRESS/SS is calculated where SS (sum of squares) is the residual sum of squares of the previous component.

According to UMETRICS (2005) a good QSAR model has R^2 and Q^2 values approximate to 0.78 respectively 0.65 or higher. The model did not reach this standard but were considered good enough to be used for this study 's particular purpose.

Performance

The performance of the model was evaluated using the root mean squared error of estimation (RMSEE). RMSEE specifies the average error in the predictions (absolute terms) (Eq. 4). No external validation was performed. The statistics of the model is presented in table 4.2.

$$\text{RMSEE} = \frac{\sum_i (y_{obs} - y_{pred})^2}{n-1-A} \quad \text{(Eq. 4)}$$

Table 4.2 Statistical performance of the model selected for the calculation of *P. Subcapitata* EC50. N: Number of observations. R2X(cum): Cumulative SS of the entire X explained by all extracted components. R2Y(cum): The cumulative SS of all the y-variables explained by the extracted components. Q2(Cum): The cumulative Q2 for all the y-variables for the extracted components. RMSEE: Root Mean Square Error of the Estimation (the fit) for observations in the workset (UMETRICS, 2005).

Type	N	R2X(cum)	R2Y(cum)	Q2(cum)	RMSEE
PLS	78	0.78	0.653	0.496	0.636

4.3 Analysis

4.3.1 Data availability

Out of the 159 plastic additives of concern data availability of toxicity of the three phyla resulted in that CFs could be derived for in total 97 substances; 80 for dataset 1 (QSAR predicted data for 2 phyla, fish and daphnia), 44 for dataset 2 (QSAR predicted data for all 3 phyla, fish, daphnia and algae) and 39 for dataset 3 (experimental data for all 3 phyla). Despite the efforts to collect more data, there were still large data gaps on the substances of concern (Figure 4.1). All derived CFs are listed in Appendix B.

For dataset 1 and 2 there was an overlap of 38 substances out of which 5 had isomeric structures resulting in totally 44 corresponding structures on both sets. SMILES strings of the structural isomers are listed in Appendix E.

Dataset 3 contained 19 substances that were also in dataset 1 and 2. Experimental CFs for 3 additional substances (84-75-3, 128-37-0, 68515-51-5) were derived using recommended effect data from USEtox's organic database resulting in totally 22 overlapping substance CFs over all three sets.

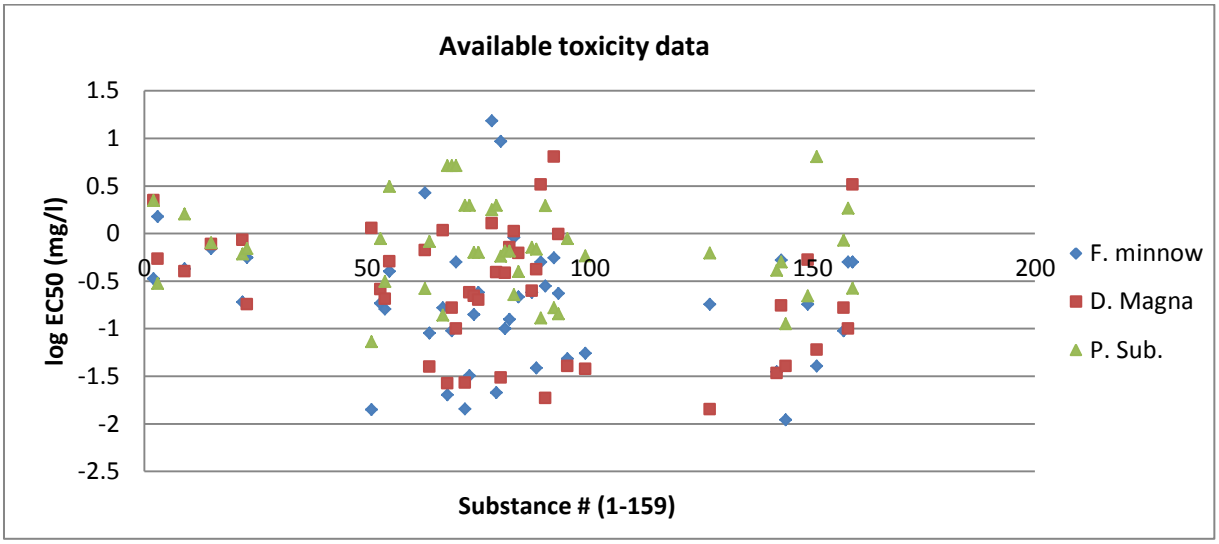

Figure 4.1 Available effect data of all species. Values on EC50 (mg/l) are plotted on a log scale and substances are distributed on the x-axis. Blank areas on the x-axis represent substances with data gaps and where no CF could be derived.

4.3.2 Substances of most concern

CAS numbers for the 10 highest ranked substances according to CFs representing impacts for emissions to continental fresh water (PAF.m^3.day.kg^{-1}) are shown in Table 4.3. No substances are found on all three lists, however, 2,2'-((1-methylethylidene)bis(4,1-phenyleneoxymethylene))bis-oxiran, 2-(3-hydroxyquinolin-2-yl)-1h-indene-1,3(2h)-dione and Tritolyl phosphate match in dataset 1 and 2 (bolded in black) and 1,3-Dichloro-2-propanol phosphate (3:1) and Triphenyl phosphate in dataset 3 match the other two lists (bolded in red). The small agreement between the top ten substances is not alarming, since as is shown below, the overall ranks did not vary much between the different lists.

Table 4.3 The table displays the substances (CAS) with the top 10 highest CFs in descending order in all three datasets. Substances that occur in more than one dataset are marked bold. Total number of substances in the datasets are also included in table.

Rank	*CAS Dataset 1*	*CAS Dataset 2*	*CAS Dataset 3*
#subst. in DS	*80*	*44*	*39*
1	**13674878**	**1675543**	64359815
2	**1675543**	**7576650**	133073
3	37853591	3896115	1118463
4	30125474	26444495	133062
5	**7576650**	**1330785**	683181
6	87843	26444495 (2)	25637994
7	**1330785**	78320	131577
8	13674845	**115866**	**13674878**
9	60348609	68411461	**115866**
10	20566352	85687	118796

4.3.3 Algal impact on substance characterization factors

Including algal toxicity in the assessment of CFs resulted in lower CF compared to CFs derived based on fish and Dapnia only (independent samples t-test, p=0.02). There were high correlations between the 44 CFs derived from dataset 1 and 2 (figure 4.2a, R^2=0.921) and between the ranks of the 44 substances (Figure 4.2b, R^2= 0.844). Ranks are studies as they show the internal orders of the substances , and how they consequently would be prioritized in a screening based on CFs. This shows that the impacts of including algal data lowered the level of the CFs, but had a small impact on the relative order between substances.

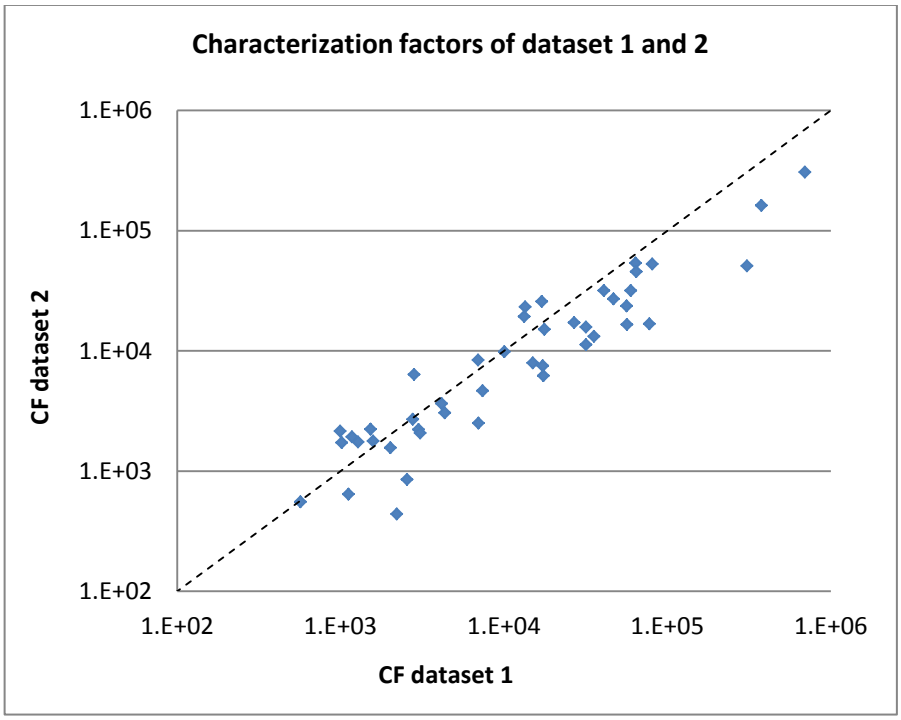

Figure 4.2a Dataset 1 and 2 are plotted against each other based on value of CFs (PAF.m3.day.kg-1) (R^2=0.921) The dashed line represents a 1:1 line.

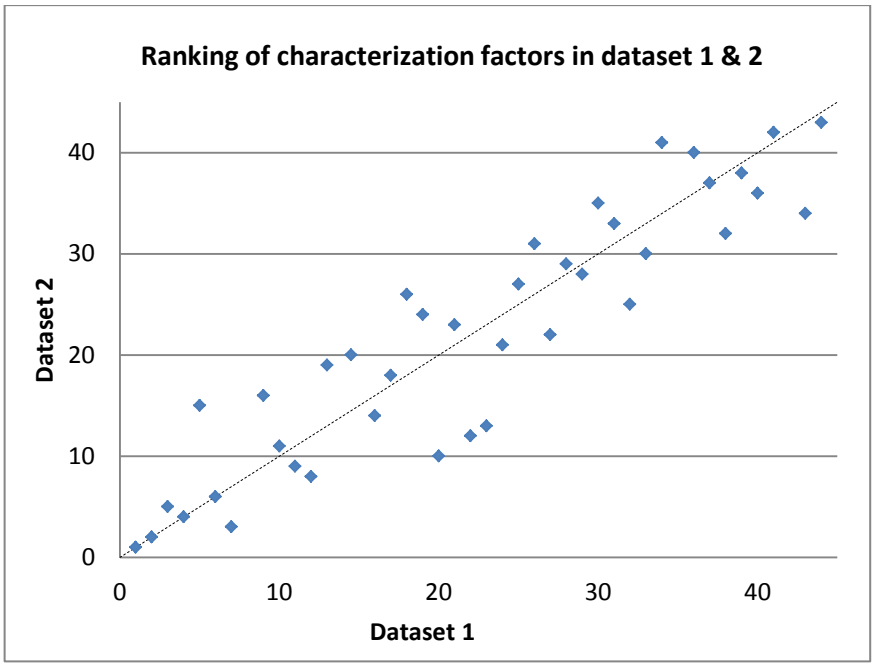

Figure 4.2b Rank of CFs for each substances in dataset 1 and 2 (R^2=0.844). The dashed line represents a 1:1 correlation. The lowest rank scores represent the highest CFs.

To further illustrate the differences in ranks following including algal toxicity, substances were color-coded into five levels where the red represents the highest ranked substances and green the lowest ranked substances based according to CFs derived from dataset 1. Including algal toxicity results in a rearrangement on some of the substances, but the majority stay within their original level (compare upper layer to bottom layer in Figure 4.3).

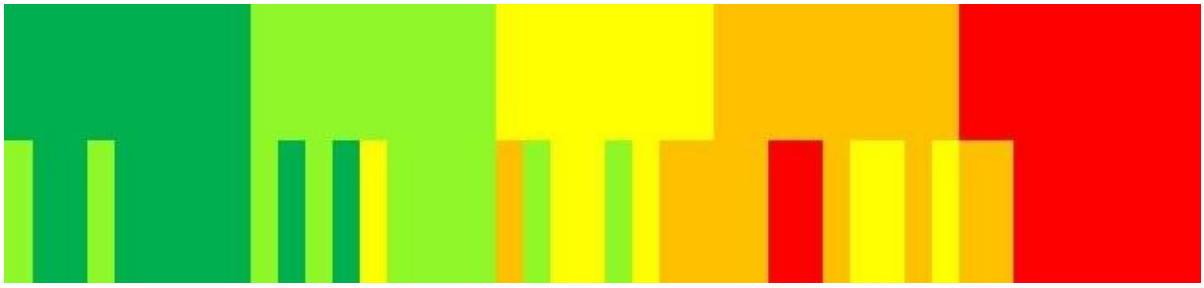

Figure 4.3 Illustration of the change in substance rank based on CF. Top layer: rank in dataset 1, Bottom layer: rank in dataset 2. Red represents the highest ranked substances in dataset 1 and dark green the lowest.

4.3.4 Comparison of effect data to investigate algal impact

Differences in CFs between the datasets depend solely on differences in effect data and therefore the variation between the species were assessed. Results show that *P. Subcapitata* in general is less sensitive to the additives and also that it exhibits smaller variation in sensitivity to the different compounds compared to *D. Magna* and *F. Minnow* (Figure 4.4).

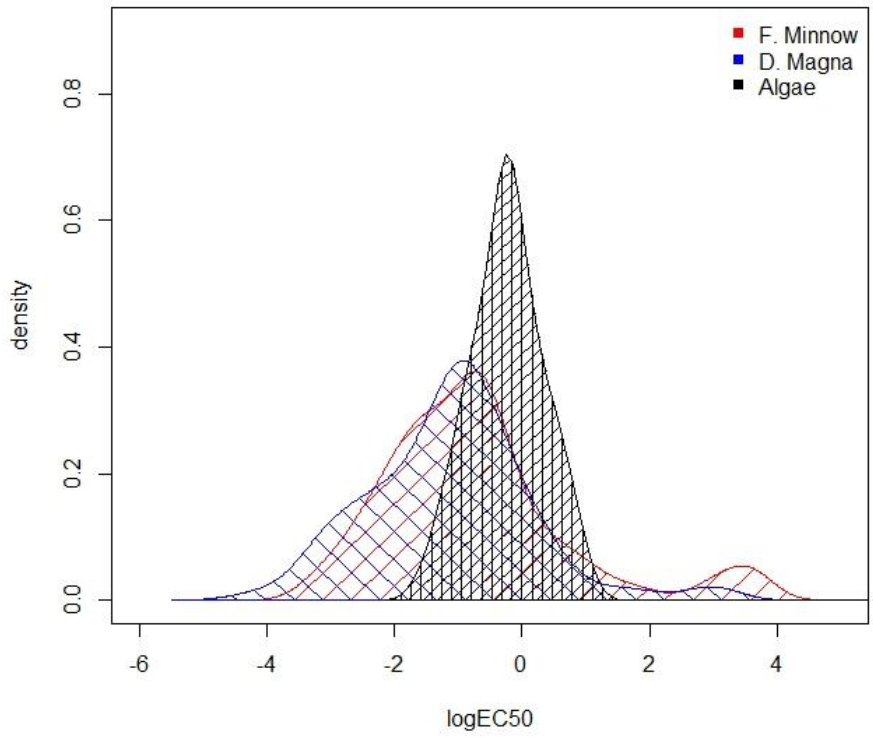

Figure 4.4 Histogram displaying log EC50 values for each species on the x-axis and proportion on the y-axis

Differences were also assessed by a Kruskal-Wallis H test using SPSS 22.0. The test showed that there was a statistically significant difference of sensitivity between the species $\chi 2(2) = 14.36$, $p = 0.001$. The null hypothesis of sensitivity being equal between the species, assuming the data points on the list is a representative random sample, could therefore be rejected. The overlap of D. Magna and F. Minnow shown in the histogram (figure 4.4) as well as the mean rank (F. Minnow: 51.33, *D. Magna*: 61.38 and *P. Sub*.: 86.80) indicates that this difference lies between algae and the two other groups.

A Mann-Whitney U test confirmed that the sensitivity (logEC50) of P. Sub. (Mdn=-0.186mg/l) were lower than sensitivity of D. Magna (Mdn=-0.594), U = 601, p = .002, r =-0.327. The sensitivity of P. Sub was also lower than the F. Minnow (Mdn=-0.727), U = 442, p < .001, r =-0.468. However no statistically significant difference could be seen between the sensitivity of D. Magna and F. Minnow, U = 826.5, p = .238, r =-0.126. This implies that the D. Magna and F. Minnow overall have similar sensitivity, and that P. Sub. in general is more tolerant to the substances considered, which explains the lower CFs for dataset 2 (Figure 4.2a).

Locking at substance specific differences between the P. Sub. and the other two species there is a slight negative correlation and a large spread between the data (Figure 4.5, R^2=0.084) implying that algae is sensitive to other substances than D. Magna and F. Minnow. The inter species variation is a likely explanation to why all substances except 4 had a different ranking when including algae.

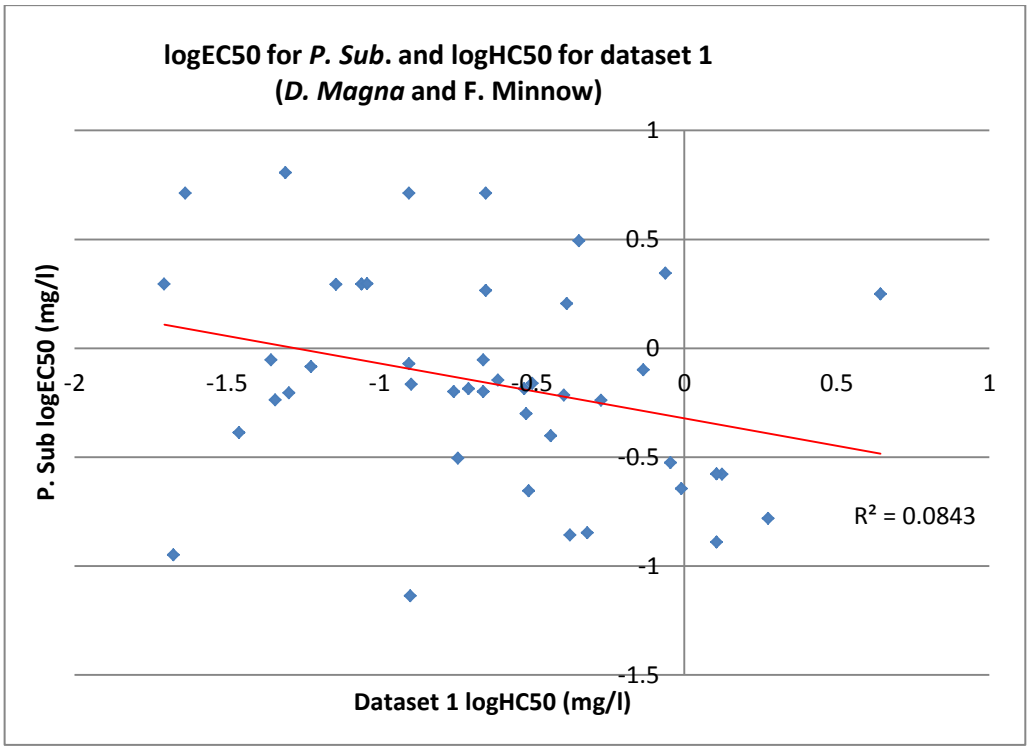

Figure 4.5 Regression plot of effect data for algae (P. Sub.) and hazard concentration for Dataset 1 (D. Magna and F. Minnow) (R^2=0.084). The red line represents the trend of the data.

The uncertainty analysis made it possible to derive 95% confidence intervals (CI) on CFs calculated based on randomized experimental algae data combined with data for D. Magna and F. Minnow (black interval in Figure 4.6). The intervals show what impact to expect by including algal toxicity without any substance specific information. Comparing these intervals to the CFs derived using the factual experimental effect data for *P. Subcapitata* (red series in Figure 4.6) showed that 88% of the substances with experimentally derived CFs were inside the 95% intervals. Thus, the CFs calculated based on factual experimental algal data (red series) are not deviating from the CFs calculated with non-substance specific information on algal toxicity. This confirms that including algal information does not change CF much since there is no strong relationship in sensitivity between algae and the two other species.

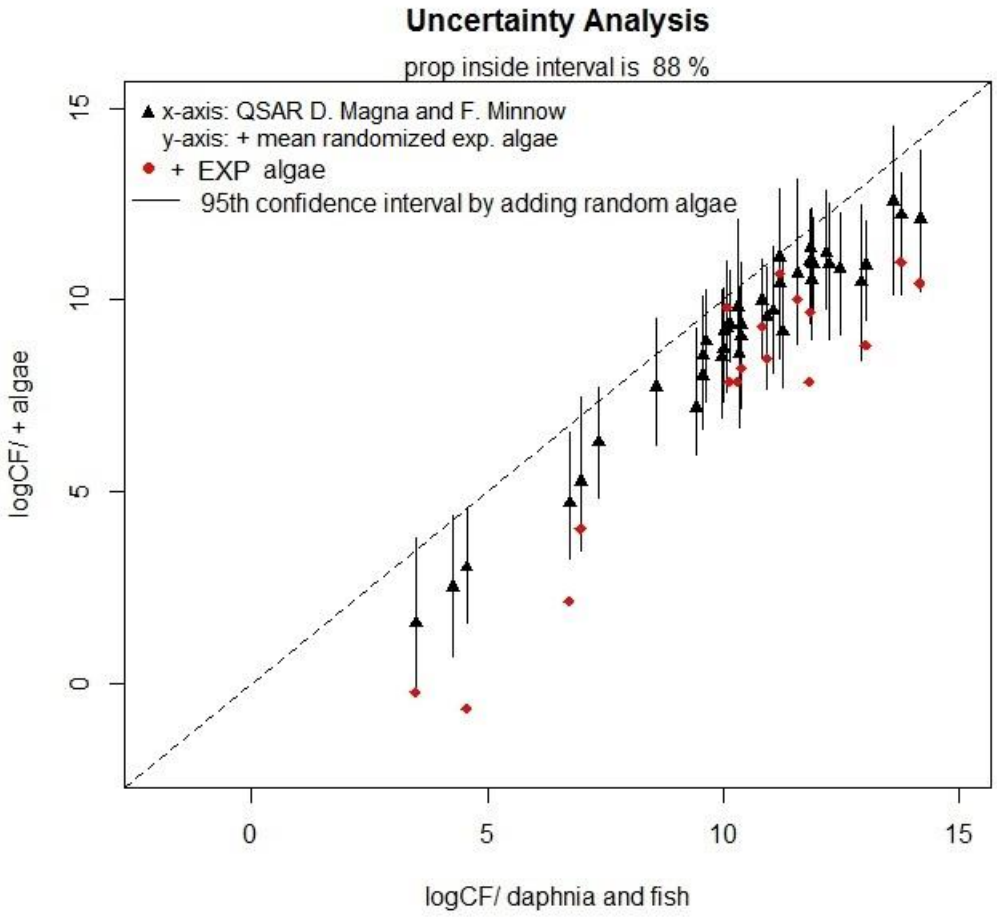

Figure 4.6 Uncertainty analysis of logCF (PAF.m3.day.kg-1) when algae is included (y-axis) versus not included (x-axis). Black interval: Interval of CFs based on dataset 1 + randomized samples of exp. data for *P. Subcapitata*. Black points: x-axis = CFs based on dataset 1, y-axis= mean of randomized samples. Red points: CFs calculated based on dataset 1 + factual experimental algal data. The dashed line represents a 1:1 correlation.

4.3.5 Differences in QSAR and experimentally based characterization factors

QSAR predicted algal data showed a relatively strong correlation with experimental algal data (R^2=0.5135) and the points are spread out on both sides of the 1:1 line indicating that there is no notable trend of over- or underestimation in the QSAR predictions compared to the experimental data (figure 4.7).

The comparison of dataset 2 and 3 (figure 4.8) on the other hand show that 7 out of 22 CFs were higher in dataset 2 than in dataset 3 implying that QSAR generally overestimates hazard compared to experimental data. The correlation between experimental and QSAR data is moderate.

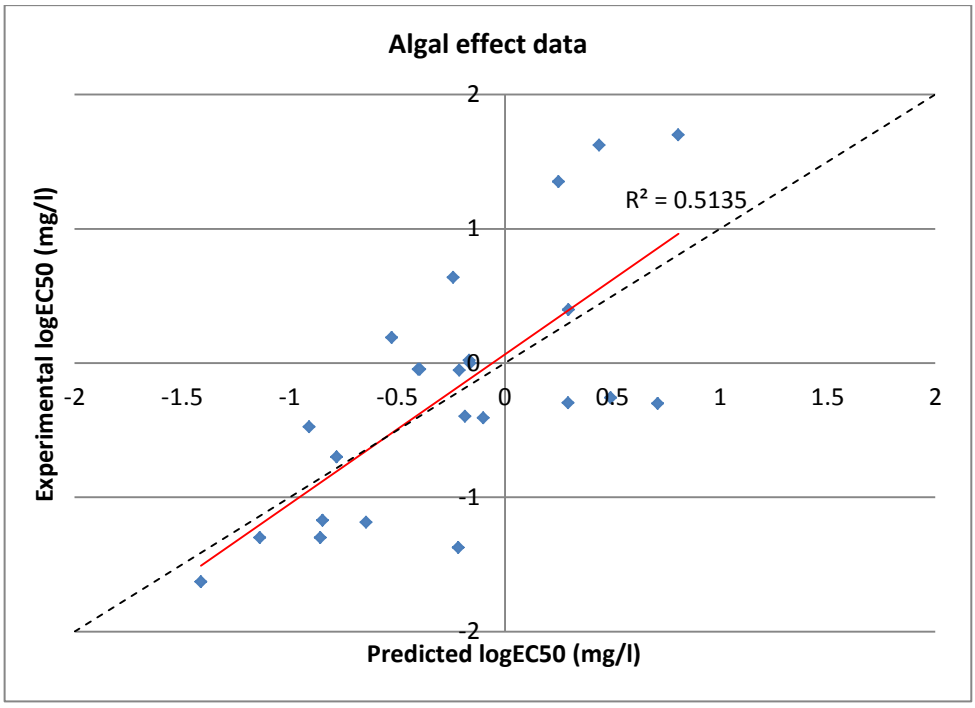

Figure 4.7 Plot of experimental versus predicted algal effect data logEC50 (mg/l). The dashed line represents a 1:1 correlation and the red line represents the trend of the data.

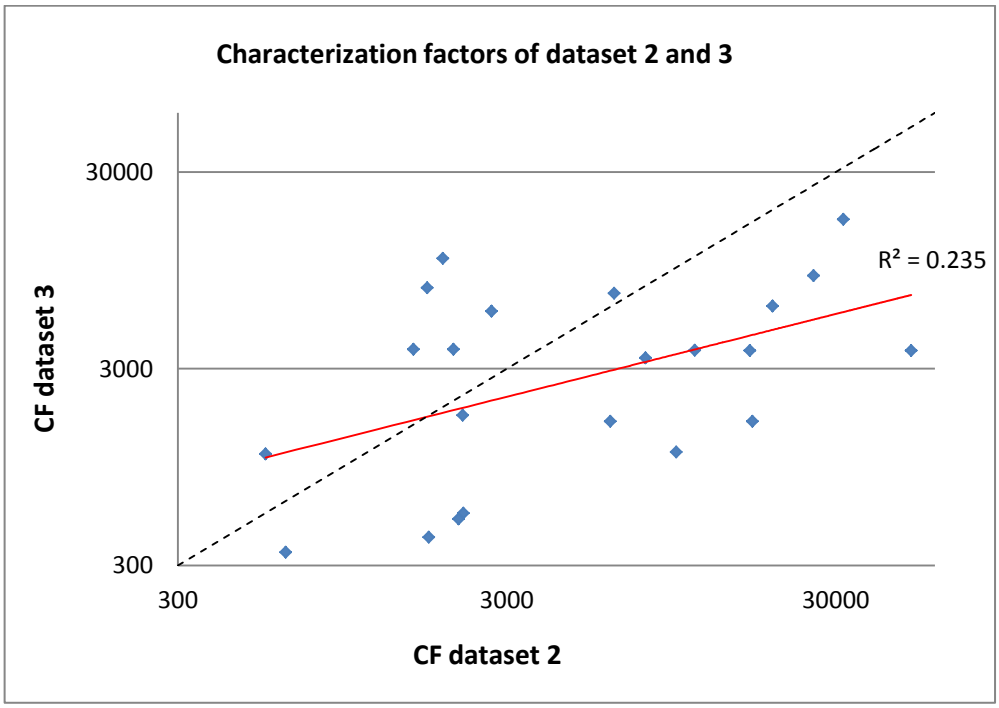

Figure 4.8 Plot of CFs for 22 substances in dataset 2 and 3. The dashed line represents a 1:1 correlation and the red line represents the trend of the data.

5. Discussion

5.1 The algal data gap and the potential of QSARs

Despite a search for existing toxicity data in some of the largest databases in the world, only 39 out of the 159 plastic additives considered had data from publically available experimental tests for algal toxicity. Further, existing QSAR models for algal toxicity could only predict toxicity for three of the substances. This highlights the large data gap for aquatic toxicity on algae, which also was shown by Payet (2004) who found that algal data represented only ca 5% of the acute aquatic toxicity data when reviewing six of the largest databases for aquatic toxicity. The scarce experimental data can partly explain the lack of algal QSARs, which have the potential to fill some of the data gaps. Here, a QSAR were constructed that could make predictions for 54 of the substances, still leaving 94 of the total 159 substances with neither experimental nor QSAR predicted values on algal toxicity. A relevant question to ask is what implications data gaps may have on chemical safety. When no data of high quality is available, are we to use default values, conservative values or not carry out any assessments at all? For example, how should the forthcoming Life Cycle Assessment on plastic additives treat the lack of data on a large part of the substances on the list? The small agreement between the three sets' top ten most hazardous substances in this study was not primarily due to differentiating data but due to data gaps leaving a large part of the substances unassessed. The need to overcome data gaps in chemical safety has led to an enhanced use of non-testing information (van Leeuwen et al., 2009), but such information is not unproblematic (see discussion below). The questionable agreement between CFs based on experimental data only and CFs based on QSAR data, shown in this study, is a potential concern, which highlights the need to address uncertainty and weaknesses in different types of data, such as testing and non-testing information, in impact assessments (Hung and Ma, 2009, Sahlin, 2013).

5.2 The impact of algal toxicity on life-cycle impact assessment of plastic additives

A question asked was; what difference it makes if algal toxicity is included or not, and why. Such questions provide important understanding of what the consequences of including and excluding different phyla in LCIAs are and for which phyla data is available. Varying sensitivity between phyla can depend on differences in biological response and the results highlight the importance of including species from various phyla to be able to account for variations in biological systems. Here, algae respond differently to the substances on the list and generally seems to be less sensitive to the substances than the other two species.

USEtox guidelines of using at least 3 phyla when calculating HC50 is based on the AMI method developed in a study by Payet (2004) where he studied how the reliability in HC50 varied depending on how many phyla were included. He found that the confidence intervals and correlation when using one, two or 3-5 phyla were [0.01; 246.93] R=0.56; [0.07; 8.12] R=0.86; [0.28; 3.88] R=0.95 respectively. Hence he concluded that using one or two phyla would be too uncertain. The difference between the species HC50 in this study highlights that the sensitivity among phyla can vary greatly and thereby supports Payet's conclusion of the importance of including toxicity data from several phyla.

The goal when conducting LCIAs is to achieve as high physiological variability as possible; ideally as many species and phyla as possible would be included (Henderson et al., 2011). When available effect data was assessed by Payet (2004) it was found that chordate and arthropod data represented

56% and 30% of the data respectively. Since data for chordate and arthropods are so dominating, including as many species as possible and calculating HC50 on a species-specific basis could lead to an under representation of effect data of phyla for which data are more scarce. The results from this study suggest that underrepresentation of algae would generally lead to higher HC50 and thereby more conservative CFs.

Larsen and Hauschild (2007b) assessed ecotoxicity effect indicators (EEI) for USEtox and concluded that to obtain a more representative CF; instead of calculating the HC50 based on species-specific EC50 the "GM-troph" method should be used. The method involves calculating the geometric mean of three different levels; species, genus and trophic level where totally three trophic levels are included; algae, invertebrates (crustaceans) and fish. This would put equal weight on each trophic level and thereby reduce the chance of biased data due to unequal representation of trophic levels. In a another study Larsen and Hauschild (2007a) suggested that to improve the PAF-based effect indicators; species and phyla included in the LCIA studies should not be randomly collected but be selected more carefully based on possibly affected ecosystems.

The comparison between phyla made in this study supports the suggestions of Larsen and Hauschild by showing that P. Sub. exhibits sensitivity to other substances than the chordate F. Minnow and the arthropod *D. Magna*. Despite this the rank of the substances were not significantly affected by including algae. A probable explanation is firstly; the small variation in the algal effect data, secondly; that two out of three data points on which HC50 was based were the same as well as all physiochemical data in dataset 1 and 2 making the influence of the algal factor minor. The algal data did however significantly alter the absolute values of the CFs with an average of 20% decrease in substance CF when algal effect data were included.

The intended LCIA will use the actual CFs to determine the relative impact of different chemicals. The results in this study suggest that including algae would not give a significantly different outcome in a relative assessment such as a LCIA. Mixing assessments with and without algal toxicity may however lead to differences between substances that depend on differences in underlying phyla and not due to actual differences in their impact.

This highlights the importance and difficulties of including algal effect data in LCIAs for this chemical group and stresses the need for further studies of algal toxicity to increase the availability of experimental data as well as new QSARs covering a wider domain of applicability.

5.3 Differences in QSAR and experimentally based characterization factors

The comparison of CFs based on experimental and QSAR data were based on a small dataset (22 substances) due to a minor overlap between the sets and hence no confident conclusions can be drawn. Yet, that about 70% of the substances got higher CFs when QSAR data was used indicates that using QSAR predicted data mostly leads to more conservative CFs. The low correlation between QSAR and experimentally based CFs (R^2=0.235) could either be due to low quality predictions e.g. that predictions are made on the edge of the models AD or that the experimental data for some reason is not representative e.g. overrepresentation of species with more extreme sensitivity/tolerance to the specific substances than *D. Magna*, F. Minnow and *P. Subcapitata.*

Few studies are available where QSAR has been compared to experimental data in an LCIA context. However, one study was found that assessed differences in experimental versus predicted and extrapolated chronic toxicity data for four structured analogues of chlorinated anilines (Dom et al., 2012). Here it was found that the QSARs and set ACRs could not account for the inter-substance and

inter-species variations. Hence, ECOSAR produced predictions that sometimes overestimated and sometimes underestimated toxicity, similarly to what the results showed in this study, and the ACR determined in *D. Magna* were greatly differentiating between the 4 analogues. Another interesting finding was that the fundamental relationship on which most QSARs are based; higher log K_{OW} equals higher toxicity, was not applicable for these substances, but an opposite trend was seen.

The study highlights the complexity in predicting biological responses and problematics in using fixed ACRs (discussed in section 5.5 and 5.4 respectively).

In this study Dihexyl phthalate (84-75-3) and Diisooctyl phthalate (27554-26-3) exhibited the largest difference in QSAR and experimental CFs. Dihexyl phthalate is one of the three substances for which the effect data was imported from USEtox database and therefore it is not possible to evaluate the source. For Diisooctyl phthalate the experimental effect data were based on five species of chordate, three species of arthropod and one species of algae with EC50's ranging between 0.13-0.55mg/l. The predicted effect data were 1.82, 2.1 and 0.45 mg/l for chordate, arthropod and algae respectively. Apparently, for this substance only the QSAR constructed in this study made predictions similar to the experimental data while TEST's predictions resulted in effect concentrations 3 or 4 times higher. In this case the experimental EC50 was based on can be considered rather reliable since it was based which all exhibited similar sensitivity. It is therefore likely that the predictions derived from TEST are uncertain for this particular substance.

5.4 Handling of information in LCIA

Consistency in the information on which a LCIA is based is essential to produce an accurate assessment. The use of different types of data is a common problem, such as the use of acute or chronic data. In LCA studies chronic data are typically preferred as it is considered more reliable. Using acute data for LCA studies has low environmental relevance since the in vivo exposure to ecosystems could realistically only be chronic (Larsen and Hauschild, 2007a). With acute data there is also a chance that the steady state between test species and the test medium was not reached (Payet, 2004).

Even though USEtox recommends that preference should be given to chronic data, in this study chronic data extrapolated from acute data was used exclusively. The motive was mainly that by treating the experimental data the same as the QSAR data the CFs would be more analogous. Another reason was the lack of chronic data. When Payet (2004) assessed the availability of effect data in the largest available databases for aquatic toxicity he found 109840 values for acute EC50 versus 9313 sub-chronic and chronic values. He also found that acute and chronic data are highly correlated (R^2=0.93). Thus, in the same way as basing CF assessments with and without algal toxicity, the results by Payet (2004) suggests that even though the absolute value of substance CFs might have been different if chronic data had been used, using acute data did not necessarily affect the rank of the substances. Finally, there are uncertainties related to factors used to extrapolate from acute to chronic, since they differ depending on e.g. what type of substance is being assessed and if it has a specific TMoA (toxic mode of action) or not (Larsen and Hauschild, 2006). This was demonstrated in the study by Dom et al. (2012) mentioned in section 5.3. Therefore even though the experimental CFs calculated in this study are recommended according to USEtox guidelines, using chronic data for substances where such data were available likely would have improved the quality of the CFs.

Another source of error that could influence the LCIA is the physiochemical data which here were derived using EPIsuite, which is a collection of estimation programs for physiochemical property and

environmental fate data. Both EPA and USEtox developers recommends that if measured values are available these should be used instead (USEtoxTeam, 2014).

In this study solely EPIsuite data was used for all physiochemical data as well as BAF. This might have increased the uncertainty in the CFs that was derived. Yet, the foremost purpose of this study was to evaluate the applicability of QSAR data in USEtox and what impact algal data has on the derived CFs. Similarly to using acute data instead of chronic; using predicted physiochemical data has not affected the results of this evaluation since the data was handled the same for all three sets.

5.5 Reliability in experimental and QSAR information

The experimental data used in the study was mainly obtained from EPA, QSAR toolbox and Aiida containing data from predominantly peer reviewed scientific literature. Yet, for several substances the effect data were scarce and hence the HC50 was calculated based on only a few data points which increase the chance of misrepresenting HC50s.

Due to the short time limit of this study the original sources from which the experimental data originated were not examined in detail. In the short summaries provided by the Aiida database the data was rated depending on its reliability from 1-4 where 1 equals “reliable without restrictions”. Many of the data points were marked as 2: “reliable with restrictions” or 3: “not reliable”. Sources marked 3 or lower were not included in this study.

A typical issue was that the effect concentration exceeded the maximum solubility of the substances (this was the case for many effect data on algae). Hence, no EC50 concentration could be calculated but the > (greater-than signs) were used combined with the highest obtained concentration. This is a common issue when testing substances that are poorly soluble. In this paper the experimental details such as how test solutions were prepared and how undissolved test solution was handled was not taken into account when collecting the data although it is well known that different testing methodologies can produce significantly different results. According to Weyman et al. (2012) if tested endpoints are above a substance solubility limit and no effect can be observed the effect concentration should be considered to be the solubility limit. Weyman et al. (2012) also suggested that if a substance exerts no toxicity at its solubility limit it could be an indication that the substance has a water solubility below the $ETNC_{aq}$ (aquatic exposure threshold of no concern). Hence, it would present a low environmental risk and could be down prioritized or excluded from the study in an early phase. These occurrences were not accounted for when collecting data in this paper.

Another source of uncertainty in the experimental data is that the reported EC50s differed between the sources. Sometimes this difference could be as great as up to a factor of 100. In these cases the data that had the highest reliability according to the Aiida rating system was used. For data obtained at ECOTOX or QSAR toolbox the value most similar to other analogous species were used.

The reliability of information becomes more complex to evaluate for QSAR prediction compared to experimental information. There are several extra sources of uncertainty to consider. Firstly, predictions from a model always come with an associated model error. Still, QSAR predictions are often given as point estimates, also in this study. Generally, QSAR models predicting toxicity data are associated with larger uncertainty than models predicting e.g. physiochemical properties. This is due to the larger variations associated with biological endpoints since there are numerous of factors that influence how organisms reacts (Sahlin, 2014).There are several ways to limit these variations; one way is to use consistent training data that ideally originates from the same study (Rodgers et al., 2011). Secondly, the extent to which a substance falls inside a model’s domain of applicability

determines the quality of the predictions. Therefore relevant measures should be used in the evaluation phase to account for the reliability in the predictions. The data obtained from TEST is considered quite reliable since the software does not make predictions for substances outside the models' applicability domains.

Finally, the quality of a QSAR depend on the quality of the training data. In TEST all experimental data that were used to train the models for *D. Magna* and F. Minnow was obtained from EPA's ECOTOX database which consists of mainly peer-reviewed data (EPA, 2014). However, which models were used for which substances and how applicable and accurate these models were has not been evaluated in this study.

The data used to construct the algal model was collected from reliable sources e.g. EPA, ECHA and Aiida that contain mostly peer-reviewed papers but originated from a vast amount of different studies. It is possible that the test methodology differed in these studies which would increase the uncertainty of the model predictions. Furthermore, it is essential to have a clear biological endpoint. For this model a single species of algae was used (*P. Subcapitata*) but due to scarce ecotoxicological data on the assessed substances the observation duration ranged between 72-96h, instead of only one fixed time. This might have contributed to some variation in the data underlying the model.

5.6 Environmental relevance

Plastic additives are a group of chemicals that have often been disregarded in LCA literature (Van der Voet, 2013). In a study by Westerdahl et al. (2010) emissions of plastic additives from Swedish societal material stock was assessed where results showed that 2% corresponding to ca 50 000 tonnes were emitted to the environment each year. In this paper the impacts of a group of plastic additives that was identified for constituting a large part of the Swedish emissions in previous study by Westerdahl et al. (2010) was assessed. To accurately assess impacts and minimize the risk of damaging ecosystems there is a need to include a high physiological variability in species when conducting LCIAs (Henderson, 2011). However, the review of ecotoxicological data in this study confirms what other studies has concluded that that there is a large gap in the toxicity data for aquatic organisms and that algal data is particularly scarce (Larsen and Hauschild, 2006, Payet, 2004). Therefore, another aim of the study was to assess the possibility of using QSAR-predicted toxicity data where experimental data is missing.

This study has resulted in a compilation of toxicity data for plastic additives for species from three aquatic phyla; chordate, arthropod and algae. A QSAR-model has also been constructed based on the experimental data that was gathered. The CFs that has been derived will be used for further studies of environmental impact where they combined with the emission loads derived by Westerdahl et al. (2010) can be used to calculate total impact scores (IS) for the additives. The QSAR model that has been constructed and the data that has been compiled can also be used in future studies of plastic additives e.g. in LCIAs.

6. Conclusions & recommendations

- There is a large data gap in algal toxicity for the 159 plastic additives of concern.
- It was possible to construct a QSAR model for algal toxicity that was applicable for about 1/3 of the substance on the list.
- CFs have been derived to identify the substances of most concern, but are limited to substances for which experimental or QSAR information are available.

- The sensitivity of the species varied both between phyla and between substances implying that which phyla is included is crucial for outcome of the impact assessment.
- Algae generally exhibited lower sensitivity to the substances on the list. When algae was included in the HC50 the CFs decreased by an average of 20%. Underrepresentation of algae in LCIA studies of similar substances might therefore lead to more conservative CFs.
- The inclusion of algal effect data in logHC50 did alter the internal rank of the substances based on CFs although no extreme alterations were seen. This is most likely due to the low variation in sensitivity that algae expressed to the different substances compared to the other two species. When assessing plastic additives including algal data is therefore of greater importance if the purpose is to derive valid CFs than if the purpose is to assess relative toxicity.
- Assessments based on experimental data are different to those based on QSAR predictions, but the difference may not be larger than the errors due to other sources of uncertainty such as quality in data, ambiguous endpoints, different representation of certain species and phyla, extrapolations between acute and chronic tests, model errors, and limited applicability of models.
- More tests of algal toxicity to plastic additives would not only lead to a more representative species distribution in online databases, it would also improve the possibilities to construct QSAR-models for algal toxicity.

7. Future work

- In this study the minimum diversity; three phyla to produce a recommended result was used for all datasets. There are several important freshwater phyla that were not included in the study e.g. *Protozoans, Rotifera, Cnidarians, Platyhelminthes* and *Mollusca*. To produce more accurate CFs and also to be able to make more valid comparisons between QSAR and experimental data future studies could involve a higher diversity of species.
- One of the conclusions in this study is that algal data does affect the CFs significantly and also that different phyla are sensitive to different substances. Therefore, it would be interesting to evaluate if the experimental CFs would differ if the logHC50s were calculated based on a phyla level where average of each phyla would be derived before calculating the mean e.g. using the GM-troph method (Larsen and Hauschild, 2006). This would prevent overrepresentation of e.g. arthropod and chordate effect data.
- The substances that were assessed in this paper were selected based on lists that Swedish research program ChEmiTecs assembled when assessing emissions of additives from plastics in the Swedish societal material stock. Future work might include combining CFs derived in this study with mass emissions for calculation of impact scores (IS) which is a more environmentally relevant measure. An interesting approach would be to examine if the substances are ranked similarly when emissions are considered.
- Due to scarce data the comparison between CFs derived using experimental versus QSAR data were based on only 22 substances. There is a need for more comprehensive studies to draw any certain conclusions about the compatibility of QSAR in LCIA.
- Further sensitivity analyses of the parameters in USEtox are needed to determine the importance of accuracy in different input data.
- In this study physiochemical data was predicted using EPA's EPIsuite software. This might have affected the CFs. A proposal for additional studies is therefore to address the uncertainty of using QSAR predicted physiochemical data as opposed to experimental.

8. Acknowledgements

I wish to acknowledge my supervisors during this project for their invaluable assistance and guidance;

A special thanks to Ullrika Sahlin at the Centre for Environmental and Climate Research (CEC), Lund University, for all help you have offered both by providing constructive feedback and comments on my work and by assisting me during the analyses.

I would also like to express my great appreciation to my external supervisors Tomas Rydberg and Magnus Rahmberg at IVL, Swedish Environmental Research Institute.

Tomas gave me the chance to learn more about QSAR-modelling and LCIA by offering me this assignment and has contributed with support and guidance throughout this process.

Magnus has provided me with all the tools I needed for the compilation of data related to the construction of the QSAR-model. He also patiently guided me through the different steps in the QSAR modelling process and helped me to understand the basics of QSAR-modelling.

Finally I would like to thank Hanna Andersson at Chalmers University of Technology for offering help and providing me with valuable material for this this study.

9. References

ARUOJA, V. 2011. *ALGAE PSEUDOKIRCHNERIELLA SUBCAPITATA IN ENVIRONMENTAL HAZARD EVALUATION OF CHEMICALS AND SYNTHETIC NANOPARTICLES.* PhD, Estonian University of life sciences.

BARE, J. C. 2006. Risk assessment and Life-Cycle Impact Assessment (LCIA) for human health cancerous and noncancerous emissions: Integrated and complementary with consistency within the USEPA. *HUMAN AND ECOLOGICAL RISK ASSESSMENT,* 12**,** 17.

DOM, N., KNAPEN, D. & BLUST, R. 2012. Assessment of aquatic experimental versus predicted and extrapolated chronic toxicity data of four structural analogues. *Chemosphere,* 86**,** 56-64.

EPA. 2012. *PBT Profiler* [Online]. Available: http://www.epa.gov/oppt/sf/tools/pbtprofiler.htm [Accessed 04-22 2014].

EPA. 2014. *ECOTOX User Guide: ECOTOXicology Database System. Version 4.0.* [Online]. Available: http:/www.epa.gov/ecotox/.

ERIKSSON, L., JAWORSKA, J., WORTH, A. P., CRONIN, M. T. D., MCDOWELL, R. M. & GRAMATICA, P. 2003. Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification- and Regression-Based QSARs. *Environmental Health Perspectives,* 111**,** 1361-1375.

FURUSJÖ, E., SVENSON, A., RAHMBERG, M. & M., A. 2005. The importance of outlier detection and training set selection for reliable environmental QSAR predictions. *Chemosphere,* 63**,** 99-108.

HAUSCHILD, M. 2007. GM-troph: A Low Data Demand Ecotoxicity Effect Indicator for Use in LCIA (13+3 pp). *The International Journal of Life Cycle Assessment,* 12**,** 79-91.

HAUSCHILD, M. Z. 2005. Assessing Environmental Impacts in a Life-Cycle Perspective. *Environmental Science & Technology,* 39**,** 81A-88A.

HENDERSON, A., HAUSCHILD, M., MEENT, D., HUIJBREGTS, M. J., LARSEN, H., MARGNI, M., MCKONE, T., PAYET, J., ROSENBAUM, R. & JOLLIET, O. 2011. USEtox fate and ecotoxicity factors for comparative assessment of toxic emissions in life cycle analysis: sensitivity to key chemical properties. *The International Journal of Life Cycle Assessment,* 16**,** 701-709.

HENDERSON, A. D., HAUSCHILD, M.Z., VAN DE MEENT, D., HUIJBREGTS, M.A.J., LARSEN, H.F., MARGNI, M., MCKONE, T.E., PAYET, J., ROSENBAUM, R.K., JOLLIET, O., 2011. USEtox fate and ecotoxicity factors for comparative assessment of toxic emissions in life cycle analysis: sensitivity to key chemical properties. . *The International Journal of Life Cycle Assessment 16***,** 701-709.

HERTWICH, E., MCKONE, T. & PEASE, W. 1999. Parameter uncertainty and variability in evaluative fate and exposure models. *Risk Analysis,* 19**,** 12.

HUNG, M.-L. & MA, H.-W. 2009. Quantifying system uncertainty of life cycle assessment based on Monte Carlo simulation. *The International Journal of Life Cycle Assessment,* 14**,** 19-27.

JOLLIET, O. & MCKONE, T. 2011. Rapid Exposure-Based Prioritization of Environmental Chemicals using USEtox. EPA.

LARSEN, H. F. & HAUSCHILD, M. 2006. GM-troph – A low data demand ecotoxicity effect indicator for use in LCIA. *Int J LCA.*

LARSEN, H. F. & HAUSCHILD, M. 2007a. Evaluation of ecotoxicity effect indicators for use in LCIA. *International Journal of Life Cycle Assessment,* 12**,** 24-33.

LARSEN, H. F. & HAUSCHILD, M. 2007b. GM-troph: A Low Data Demand Ecotoxicity Effect Indicator for Use in LCIA (13+3 pp). *The International Journal of Life Cycle Assessment,* 12**,** 79-91.

LINDGREN, F., NOUWEN, J., LOONEN, H., WORTH, A., HANSEN, B. & KARCHER, W. 1996. Environmental Modelling Based on a Structural Fragments Approach. *Indoor and Built Environment,* 5**,** 334-340.

LYNN, D. 2011. *The Ciliated Protozoa: Characterization, Classification, and Guide to the Literature (3 ed.)*, Springer.
MARK HUIJBREGTS, M. H., OLIVIER JOLLIET, MANUELE MARGNI, & TOM MCKONE, R. K. R., DIK VAN DE MEENT. 2010. *USEtoxTM User manual* [Online]. Available: http://www.usetox.org/sites/default/files/support-tutorials/user_manual_usetox.pdf 1.01].
MARK HUIJBREGTS, M. M., DIK VAN DE MEENT, OLIVIER JOLLIET, RALPH K. ROSENBAUM, TOM MCKONE, MICHAEL HAUSCHILD 2010. USEtoxTM Chemical-specific database: organics 1.01.
MILAN, C., ESPOSITO, A., CASSANO, A., GONELLA DIAZA, R., BENFENATI E. 2012. *Comparison of the results of QSAR models for LD50.* Istituto di Ricerche Farmacologiche Mario Negri.
NETZEVA, T. I., APTULA, A. O., BENFENATI, E., CRONIN, M. T., GINI, G., LESSIGIARSKA, I., MARAN, U., VRACKO, M. & SCHÜÜRMANN, G. 2005. Description of the Electronic Structure of Organic Chemicals Using Semiempirical and Ab Initio Methods for Development of Toxicological QSARs. *Journal of Chemical Information and Modelling*, 106-114.
OECD 2004. OECD PRINCIPLES FOR THE VALIDATION, FOR REGULATORY PURPOSES, OF (QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIP MODELS. *In:* OECD (ed.).
OECD 2007. Guidance Document on the Validation of (Quantitative) Structure Activity Relationship [(Q)SAR] Models.
OECD. 2014. *OECD Quantitative Structure-Activity Relationships Project [(Q)SARs]* [Online]. Available: http://www.oecd.org/env/hazard/qsar [Accessed 02-20 2014].
OECD. 2014a. *Donors to the QSAR Toolbox* [Online]. Available: http://www.oecd.org/chemicalsafety/risk-assessment/donorstotheqsartoolbox.htm [Accessed 2014-03-13.
PANT, R. V. H., G.; SCHOWANEK, D.; FEIJTEL, T.C.J.; DE KONING, A.; HAUSCHILD, M.; OLSEN, S.I.; PENNINGTON, D.W.; ROSENBAUM, R.K. 2004. Comparison between three different LCIA methods for aquatic ecotoxicity and a product environmental risk assessment: insights from a detergent case study within OMNIITOX. *Int J Life Cycle Assess,* 9, 295-306.
PATEL, N. & BOETHLING, B. 2006. Overview of EPI Suite™:Software for Chemical Property and Fate Estimation. *In:* TOXICS, U. S. E. P. A. O. O. P. P. A. (ed.). Washington, DC 20460.
PAYET, J. 2004. *Assessing toxic impacts on aquatic ecosystems in life cycle assessment (LCA).* Ph.D., Ecole Polytechnique Fédérale de Lausanne (EPFL).
PAYET, J. 2013. *AiiDA: Aquatic Impact Indicators Database* [Online]. Available: http://aiida.tools4env.com/index.
PUZYN, T., LESZCZYNSKI, J. & CRONIN, M. T. 2010. *Recent Advances in QSAR Studies [Elektronisk resurs] : Methods and Applications / edited by Tomasz Puzyn, Jerzy Leszczynski, Mark T. Cronin*, Dordrecht : Springer Netherlands, 2010.
RAHMBERG, M., ANDERSSON, H., WESTERDAHL, J., RYDBERG, T. & ANDERSSON, P. L. Towards QSAR based USEtoxTM characterisation factors for assessing ecotoxicological impacts of emissions from plastic additives. RISKCYCLE Conference: Risk-Based Management of Chemicals and Products in a Circular Economy at a Global Scale, 8-9 May 2012 Dresden. 47-52.
RODGERS, S. L., DAVIS, A. M., TOMKINSON, N. P. & VAN DE WATERBEEMD, H. 2011. Predictivity of Simulated ADME AutoQSAR Models over Time. *Molecular Informatics,* 30, 256-266.
ROSENBAUM, R., BACHMANN, T., GOLD, L., HUIJBREGTS, M. J., JOLLIET, O., JURASKE, R., KOEHLER, A., LARSEN, H., MACLEOD, M., MARGNI, M., MCKONE, T., PAYET, J., SCHUHMACHER, M., MEENT, D. & HAUSCHILD, M. 2008. USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human

toxicity and freshwater ecotoxicity in life cycle impact assessment. *The International Journal of Life Cycle Assessment,* 13, 532-546.
SAHLIN, U. 2014. E-mail correspondence. Stockholm.
SAHLIN, U. G., L.; IQBAL, M.S.; PEIJNENBURG, W. 2013. Arguments for Considering Uncertainty in QSAR Predictions in Hazard and Risk Assessments. *ATLA,* 41, 91-110.
SAIC 2006. Life cycle assessment: principles and practice. Washington, DC: U.S. Environmental Protection AgencyWashington, DC.
SCHAFER, H., HETTLER, H., FRITSCHE, U., PITZEN, G., RODERER, G. & WENZEL, A. 1994. Biotests Using Unicellular Algae and Ciliates for Predicting Long-Term Effects of Toxicants. *Ecotoxicology and Environmental Safety,* 27, 64-81.
SOCOLOF, M. L. O., J.G.; KINCAID, L.E.; GEIBIG, J.R. 2001. Desktop Computer Displays: A Life-Cycle Assessment. Washington D.C.
STEIN, R. S. 2004. *Pollution A-Z* [Online]. Encyclopedia.com. Available: http://www.encyclopedia.com/doc/1G2-3408100197.html [Accessed June 20th, 2014.
TODESCHINI, R., CONSONNI, V., MAURI, A. & PAVAN, M. (undated). *DRAGON Software,* Milan, Italy, Talete srl.
UMETRICS 2005. User's guide to SIMCA-P Version 11.0.
USETOX. 2013. *QSAR-models in USEtox* [Online]. usetox.org: USEtox Team. Available: http://www.usetox.org/forums/uncertainty-usetox/qsar-models-usetox [Accessed 2014-05-15.
USETOXTEAM. 2014. *QSAR-models in USEtox* [Online]. Available: http://www.usetox.org/forums/uncertainty-usetox/qsar-models-usetox.
WALKER, J. D., CARLSEN, L., HULZEBOS, E. & SIMON-HETTICH, B. 2002. Global government applications of analogues, SARs and QSARs to predict aquatic toxicity, chemical or physical properties, environmental fate parameters and health effects of organic chemicals. *SAR And QSAR In Environmental Research,* 13, 607-616.
WALKER, J. D. J., J.;COMBER, M. H. I.;SCHULTZ, T. W.; DEARDEN, J. C. 2003. Guidelines for developing and using quantitative structure-activity relationships. *Environmental Toxicology and Chemistry,* 22, 1653-1665.
VAN DER VOET, E. V. O., L.;RYDBERG, T.; WESTERDAHL, J.;LARSEN, H.F. 2013. Life Cycle Assessment of Additives: Methodology and Data. *In:* BILITEWSKI, B., DARBRA, R. M. & BARCELÓ, D. (eds.) *Global Risk-Based Management of Chemical Additives II.* Springer Berlin Heidelberg.
VAN LEEUWEN, K., SCHULTZ, T. W., HENRY, T., DIDERICH, B. & VEITH, G. D. 2009. Using chemical categories to fill data gaps in hazard assessment. *SAR and QSAR in Environmental Research,* 20, 207-220.
WESTERDAHL, J., ANDERSSON, P., FUHRMAN, F., HAGLUND, P., HALLBERG, E., HOLMGREN, T., MOLANDER, S., RYDBERG, T., TIVANDER, J. & A, Ö. 2010. National inventory of emissions of additives from plastic materials. P4-D4. .
WEYMAN, G. S., RUFLI, H., WELTJE, L., SALINAS, E. R. & M., H. 2012. AQUATIC TOXICITY TESTS WITH SUBSTANCES THAT ARE POORLY SOLUBLE IN WATER AND CONSEQUENCES FOR ENVIRONMENTAL RISK ASSESSMENT. *Environmental Toxicology and Chemistry*.

10. Appendix

Appendix A - Extracting data

Physiochemical properties

When deriving data from EPIsuite the following tools and methods were used:
Kowwin: Kow - atom/fragment contribution method (the molecular structure is divided into fragments or larger groups)
Kocwin: Koc (L/kg) MCI (Molecular connectivity index)
Henrywin: Henry's constant - Bond Estimation method
Mpbpwin: Vapour pressure – Modified grain estimation method
Waternt**:** Solubility (mg/L 25 deg. C) - regression calculations of Kow and melting point
Aopwin: kdegA (degradation parameter in air), OVERALL OH Rate Constant in cm^3/molecule/sec,

Degradation in sediment, soil and water

Degradation rate in air was calculated using the second order output OVERALL OH Rate Constant in (cm^3/molecule/sec) from AOPWIN. To get a first order rate constant the OVERALL OH Rate Constant was multiplied with 1.5E+06 OH/cm^3 and divided by 2 (assuming 12h effective removal per day) in accordance with the USEtox instructions (Henderson, 2011).

Degradation rates in water, soil and sediment were extracted from the BIOWIN 3 output (Ultimate Servey Model) and the assigned half-lifes were converted to degradation rate (1/s) as recommended by USEtox guidelines (Mark Huijbregts, 2010). The division factors 1:2:9 were used to extrapolate degradation rate in water, soil and sediment respectively as proposed by Patel and Boethling (2006).

Calculating logHC50

Instructions below were taken from the USEtox User's manual (Mark Huijbregts and Tom McKone, 2010):

1. Gather experimental or estimated EC50 data for the chemical of interest;
2. Specify for every EC50-value whether it is chronic or acute exposure;
3. Calculate the geometric mean chronic or acute EC50 (mg/l) for every individual species (this can e.g. be done with the function =GEOMEAN() in Excel).
4. In case of acute EC50-data, derive the chronic-equivalent EC50 per species by dividing by a factor of 2 (acute-to-chronic extrapolation factor)
5. Take the log of the geometric mean EC50s and calculate the average of the log-values. This average equals the logHC50 (log mg/l).
6. Implement this value in column 20 of the sheet "Substance data" of USEtox.xls.
7. Always be careful with the units

Appendix B – Substance CF

All calculated CFs are documented in the table below.

#	CAS	Name	Ecotox. Charact. factor [PAF.m3.day.kg-1] Em.fr.waterC Dataset 1 (QSAR)	Dataset 2 (QSAR)	Dataset 3 (Exp)
1	123284	Dilauryl thiodipropionate	2255.647422	n/a	n/a
2	128370	2,6-Di-tert-butyl-4-methylphenol	3008.839163	2202.910509	1744.370658
3	80057'	Bisphenol A	13327.37863	19272.62252	6256.043329
4	2082793	Octadecyl 3-(3,5-di-tert-butyl-4-hydroxyphenyl)propionate	86.31066734	n/a	0.052312318
7	31570044	Tris(2,4-ditert-butylphenyl) phosphite	n/a	n/a	0.000151156
9	4130421	2,6-Ditert-butyl-4-ethylphenol	26956.23594	17146.04205	n/a
11	6683198	Pentaerythritol tetrakis(3-(3,5-di-tert-butyl-4-hydroxyphenyl)propionate)	n/a	n/a	8.23131E-11
12	693367	Distearyl thiodipropionate	1.405886714	n/a	0.001038011
13	108781	Melamine	365.207758	n/a	26.12763238
14	118796	2,4,6-Tribromophenol	74580.39241	n/a	15371.50655
15	103231	Bis(2-ethylhexyl) adipate	568.8932894	553.9295941	1106.928555
16	21850442	Tetrabromobisphenol A bis(dibromopropyl ether)	40528.45286	n/a	n/a
17	32534819	Pentabromodiphenyl ether	22059.57922	n/a	n/a
21	615587	2,4-Dibromophenol	51198.22095	n/a	n/a
22	103242	AZELAIC ACID DI(2-ETHYLHEXYL) ESTER	4169.091516	3635.553905	n/a
23	115866	Triphenyl phosphate	41045.83211	31673.39112	17227.68843
29	30125474	3,4,5,6-Tetrachloro-N-[2-(4,5,6,7-tetrachloro-2,3-dihydro-1,3-dioxo-1H-inden-2-yl)-8-quinolyl]phthalimide	424825.1069	n/a	n/a
40	5567157	2,2'-[(3,3'-dichloro[1,1'-biphenyl]-4,4'-diyl)bis(azo)]bis[n-(4-chloro-2,5-dimethoxyphenyl)-3-oxobutyramide]	n/a	n/a	23.17124712
49	81776	14,18-anthrazinetetrone,6,15-dihydro-9	n/a	n/a	48.76689267
51	112845	cis-13-Docosenoamide	6949.533569	8349.385155	n/a
52	1843056	Octabenzone	n/a	n/a	41.57092447
53	25973551	2-(2H-Benzotriazol-2-yl)-4,6-ditertpentylphenol	7403.425031	4653.463299	n/a
54	3896115'	Bumetrizole	64136.23741	53462.84218	n/a
55	52829079	Bis(2,2,6,6-tetramethyl-4-piperidyl)sebacate	15060.67194	7919.892533	3409.904322
56	1709702	1,3,5-Trimethyl-2,4,6-tris(3,5-di-tert-butyl-4-hydroxybenzyl)benzene	7.95219E-06	n/a	n/a
58	40601761	Tris(4-tert-butyl-3-hydroxy-2,6-dimethylbenzyl) isocyanurate	194.4257585	n/a	n/a
59	25637994	Hexabromocyclododecane	n/a	n/a	117350.3584
60	25637994 (2)	Hexabromocyclododecane	20070.32284	n/a	117350.3584
63	2440224	2-(2H-Benzotriazol-2-yl)-p-cresol	13533.87645	23208.54844	n/a
64	3864991	2-(2'-Hydroxy-3',5'-di-tert-butylphenyl)-5-chlorobenzotriazole	56494.55584	23563.85591	n/a
65	63843890	Bis(1,2,2,6,6-pentamethyl-4-piperidyl) [[3,5-bis(1,1-dimethylethyl)-4-hydroxyphenyl]methyl]butylmalonate	246.0130275	n/a	n/a
66	115968	Tris(2-chloroethyl) phosphate	19685.59868	n/a	1.0151E-57
67	117817	Bis(2-ethylhexyl) phthalate	1529.983223	2216.554036	552.2827891
68	1330785	Tritolyl phosphate	308328.6521	50876.14189	3716.238034
69	1330785 (2)	Tritolyl phosphate	56830.42287	16471.72189	3716.238034
70	1330785 (3)	Tritolyl phosphate	31819.96226	11189.69329	3720.17215
71	20566352	2-(2-hydroxyethoxy)ethyl 2-hydroxypropyl 3,4,5,6-tetrabromophthalate	154340.4688	n/a	n/a
72	25155231	TRIXYLYL PHOSPHATE	77698.48979	16741.27127	1624.573337
73	25155231 (2)	TRIXYLYL PHOSPHATE	17465.43304	6189.177728	1624.573337
74	26444495	Cresyl diphenyl phosphate	80890.80502	52781.10174	n/a
75	26444495	Cresyl diphenyl phosphate	64796.85049	45524.82423	n/a
77	78400	Triethyl phosphate	4923.595185	n/a	60.3924907
78	78513	Tris(2-butoxyethyl) phosphate	1284.649648	1739.150573	418.2002481
79	119471	2,2'-Methylenebis(6-tert-butyl-4-methylphenol)	6981.037441	2502.941566	n/a
80	126738	Tributyl phosphate	10084.8192	9831.148364	1132.325665
81	26761400	Diisodecyl phthalate	3083.158494	2065.731166	3766.913872
82	26761400 (2)	Diisodecyl phthalate	2023.569795	1560.110501	3766.913872
83	27554263	Diisooctyl phthalate	1179.157304	1918.892147	10911.11788
84	28553120	Diisononyl phthalate	2770.327291	2696.204191	5891.650307
85	3319311	Trioctyl trimellitate	117.6853391	n/a	n/a
86	131577	Oxybenzone	n/a	n/a	22748.97454
87	33703081	Diisononyl adipate	4346.407765	3042.629103	n/a
88	3648202	Diundecyl phthalate	1118.372515	638.3397129	349.8429931
89	68515515	1,2-Benzenedicarboxylic acid, di-C6-10-alkyl esters	996.4809809	2139.888363	517.1081414
90	78422	Tris(2-ethylhexyl) phosphate	2557.821048	850.8155589	n/a
91	133062	1,2,3,6-Tetrahydro-N-(trichloromethylthio)phthalimide	n/a	n/a	153361.7221
92	84742	Dibutyl phthalate	2827.494675	6357.753707	7248.869072
93	85687	Butyl benzyl phthalate	17117.86533	25694.52026	8911.621274
94	13560899	1,6,7,8,9,14,15,16,17,17,18,18-Dodecachloropentacyclo[12.2.1.16,9.02,13.05,10]octadeca-7,15-diene	11035.71892	n/a	n/a
95	25973551	2-(2H-Benzotriazol-2-yl)-4,6-ditertpentylphenol	35724.55272	13156.19083	n/a
99	7576650	2-(3-hydroxyquinolin-2-yl)-1h-indene-1,3(2h)-dione	377998.9876	162023.2874	n/a
100	1118463	Butyltin trichloride	n/a	n/a	177208.0046
101	1118463	Butyltin trichloride	n/a	n/a	177208.0046
114	683181	Dibutyltin dichloride	n/a	n/a	133157.48
115	683181	Dibutyltin dichloride	n/a	n/a	133157.48
116	77587	Dibutyltin dilaurate	n/a	n/a	2230.292739
117	77587	Dibutyltin dilaurate	n/a	n/a	2230.292739
120	133073	1H-Isoindole-1,3(2H)-dione, 2-[(trichloromethyl)thio]-	n/a	n/a	243512.8889
121	26530201	2-Octyl-2H-isothiazol-3-one	n/a	n/a	215948.7598
123	64359815	4,5-Dichloro-2-octyl-isothiazolone	n/a	n/a	1380718.692
125	115775	Pentaerythritol	2.601883096	n/a	n/a
126	126589	Dipentaerythritol	15.72440591	n/a	n/a
127	25550985	DIISODECYL PHENYL PHOSPHITE	17342.56702	7506.672068	n/a
128	50704	D-Sorbitol	4.642612243	n/a	n/a
130	77996	Trimethylol propane	8.385858626	n/a	1.525866258
131	79947	Tetrabromobisphenol A	97872.72048	n/a	9729.619113
132	8013078'	Epoxidized soya bean oil	0.127495522	n/a	n/a
137	23128747	3,3'-Bis(3,5-di-tert-butyl-4-hydroxyphenyl)-N,N'-hexamethylenedipropionamide	14.5121889	n/a	n/a
138	36443682	Triethylene glycol bis(3-tert-butyl-4-hydroxy-5-methylphenyl)propionate	21764.47671	n/a	n/a
139	70321867	2-(2H-Benzotriazol-2-yl)-4,6-bis(1-methyl-1-phenylethyl)phenol	558.2697334	n/a	n/a
142	1675543	2,2'-((1-methylethylidene)bis(4,1-phenyleneoxymethylene))bis-oxiran	696363.3252	305427.0493	n/a
143	3147759	Octrizole	17747.02735	15026.20476	n/a
144	68411461	Benzenamine,N-phenyl-,reactionproductswith2,4,4-trimethylpentene	47026.21291	26936.41673	n/a
145	13674845	Phosphoric acid tris(2-chloro-1-methylethyl) ester	260129.3971	n/a	810.0422144

146	13674878	1,3-Dichloro-2-propanol phosphate (3:1)	1293241.157	n/a	17704.96039
147	3296900	2,2-Bis(bromomethyl)propane-1,3-diol	210.8037389	n/a	n/a
148	87843	1,2,3,4,5-Pentabromo-6-chlorocyclohexane	343865.8103	n/a	n/a
149	123955	N-BUTYL OCTADECANOATE	1585.509129	1772.659245	n/a
150	10081671	Bis[4-(2-phenyl-2-propyl)phenyl]amine	4785.649839	n/a	n/a
151	85609	4,4'-Butylidenebis(6-tert-butyl-3-methylphenol)	2214.248181	437.0631805	n/a
153	1843034	1,1,3-TRIS(2-METHYL-4-HYDROXY-5-TERT-BUTYLPHENYL)BUTANE	55.71427447	n/a	n/a
154	991844	2,4-Bis(octylthio)-6-(4-hydroxy-3,5-di-tert-butylanilino)-1,3,5-triazine	83.57415829	n/a	n/a
155	37853591	1,2-Bis(2,4,6-tribromophenoxy)ethane	529892.8596	n/a	n/a
156	60348609	2,2',4,4',5-PENTABROMODIPHENYL ETHER	156263.1477	n/a	n/a
157	78320	TRI-P-TOLYL PHOSPHATE	59988.14701	31716.25877	n/a
158	78308	TRI-O-CRESYL PHOSPHATE	31819.96226	15752.63613	n/a
159	84753	DIHEXYL PHTHALATE	1019.036216	1720.967781	7727.830542

Appendix C – CFs used in the analysis of dataset 1 and 2

The table contains the matching substances from dataset 1 and 2 that were used in the assessment of the effect of including algal data. The table includes predicted CFs and ranks of CFs.

			Ecotox. Charact. factor [PAF.m3.day.kg-1]			
			Em.fr.waterC		*Em.fr.waterC*	
#	*CAS*	*Name*	*Rank*	*Dataset 1 (QSAR)*	*Rank*	*Dataset 2 (QSAR)*
2	128370	2,6-Di-tert-butyl-4-methylphenol	31	3008.84	33	2202.91
3	80057'	Bisphenol A	23	13327.38	13	19272.62
9	4130421	2,6-Ditert-butyl-4-ethylphenol	16	26956.24	14	17146.04
15	103231	Bis(2-ethylhexyl) adipate	44	568.89	43	553.93
22	103242	AZELAIC ACID DI(2-ETHYLHEXYL) ESTER	29	4169.09	28	3635.55
23	115866	Triphenyl phosphate	12	41045.83	8	31673.39
51	112845	cis-13-Docosenoamide	27	6949.53	22	8349.39
53	25973551	2-(2H-Benzotriazol-2-yl)-4,6-ditertpentylphenol	25	7403.43	27	4653.46
54	3896115'	Bumetrizole	7	64136.24	3	53462.84
55	52829079	Bis(2,2,6,6-tetramethyl-4-piperidyl)sebacate	21	15060.67	23	7919.89
63	2440224	2-(2H-Benzotriazol-2-yl)-p-cresol	22	13533.88	12	23208.55
64	3864991	2-(2'-Hydroxy-3',5'-di-tert-butylphenyl)-5-chlorobenzotriazole	10	56494.56	11	23563.86
67	117817	Bis(2-ethylhexyl) phthalate	38	1529.98	32	2216.55
68	1330785	Tritolyl phosphate	3	308328.65	5	50876.14
69	1330785 (2)	Tritolyl phosphate	9	56830.42	16	16471.72
70	1330785 (3)	Tritolyl phosphate	14.5	31819.96	20	11189.69
72	25155231	TRIXYLYL PHOSPHATE	5	77698.49	15	16741.27
73	25155231 (2)	TRIXYLYL PHOSPHATE	18	17465.43	26	6189.18
74	26444495	Cresyl diphenyl phosphate	4	80890.81	4	52781.10
75	26444495 (2)	Cresyl diphenyl phosphate	6	64796.85	6	45524.82
78	78513	Tris(2-butoxyethyl) phosphate	39	1284.65	38	1739.15
79	119471	2,2'-Methylenebis(6-tert-butyl-4-methylphenol)	26	6981.04	31	2502.94
80	126738	Tributyl phosphate	24	10084.82	21	9831.15
81	26761400	Diisodecyl phthalate	30	3083.16	35	2065.73
82	26761400 (2)	Diisodecyl phthalate	36	2023.57	40	1560.11
83	27554263	Diisooctyl phthalate	40	1179.16	36	1918.89
84	28553120	Diisononyl phthalate	33	2770.33	30	2696.20
87	33703081	Diisononyl adipate	28	4346.41	29	3042.63
88	3648202	Diundecyl phthalate	41	1118.37	42	638.34
89	68515515	1,2-Benzenedicarboxylic acid, di-C6-10-alkyl esters	43	996.48	34	2139.89
90	78422	Tris(2-ethylhexyl) phosphate	34	2557.82	41	850.82
92	84742	Dibutyl phthalate	32	2827.49	25	6357.75
93	85687	Butyl benzyl phthalate	20	17117.87	10	25694.52
95	25973551 (2)	2-(2H-Benzotriazol-2-yl)-4,6-ditertpentylphenol	13	35724.55	19	13156.19
99	7576650	2-(3-hydroxyquinolin-2-yl)-1h-indene-1,3(2h)-dione	2	377998.99	2	162023.29
127	25550985	DIISODECYL PHENYL PHOSPHITE	19	17342.57	24	7506.67
142	1675543	2,2’-((1-methylethylidene)bis(4,1-phenyleneoxymethylene))bis-oxiran	1	696363.33	1	305427.05
143	3147759	Octrizole	17	17747.03	18	15026.20
144	68411461	Benzenamine,N-phenyl-,reactionproductswith2,4,4-trimethylpentene	11	47026.21	9	26936.42
149	123955	N-BUTYL OCTADECANOATE	37	1585.51	37	1772.66
151	85609	4,4'-Butylidenebis(6-tert-butyl-3-methylphenol)	35	2214.25	44	437.06
157	78320	TRI-P-TOLYL PHOSPHATE	8	59988.15	7	31716.26
158	78308	TRI-O-CRESYL PHOSPHATE	14.5	31819.96	17	15752.64
159	84753	DIHEXYL PHTHALATE	42	1019.04	39	1720.97

Appendix D – CFs used in the analysis of dataset 1, 2 and 3

The table contains the matching substances from dataset 1, 2 and 3 that were used in the assessment of differences in CFs that are based on QSAR-predicted or experimental data. The table includes predicted CFs and ranks of CFs.

			Ecotox. Charact. factor [PAF.m3.day.kg-1]					
				Em.fr.waterC		*Em.fr.waterC*		*Em.fr.waterC*
#	CAS	Name	*Rank*	*Dataset 1 (QSAR)*	*Rank*	*Dataset 2 (QSAR)*	*Rank*	*Dataset 3 (Exp.)*
2	128370	2,6-Di-tert-butyl-4-methylphenol	12	3008.839163	14	2202.910509	14	1744.370658
3	80057'	Bisphenol A	9	13327.37863	4	19272.62252	6	6256.043329
15	103231	Bis(2-ethylhexyl) adipate	22	568.8932894	22	553.9295941	18	1106.928555
23	115866	Triphenyl phosphate	4	41045.83211	2	31673.39112	1	17227.68843
55	52829079	Bis(2,2,6,6-tetramethyl-4-piperidyl)sebacate	8	15060.67194	9	7919.892533	13	3409.904322
67	117817	Bis(2-ethylhexyl) phthalate	16	1529.983223	13	2216.554036	19	552.2827891
68	1330785	Tritolyl phosphate	1	308328.6521	1	50876.14189	11.5	3716.238034
69	1330785 (2)	Tritolyl phosphate	3	56830.42287	6	16471.72189	11.5	3716.238034
70	1330785 (3)	Tritolyl phosphate	5	31819.96226	7	11189.69329	10	3720.17215
72	25155231	TRIXYLYL PHOSPHATE	2	77698.48979	5	16741.27127	15.5	1624.573337
73	25155231 (2)	TRIXYLYL PHOSPHATE	6	17465.43304	11	6189.177728	15.5	1624.573337
78	78513	Tris(2-butoxyethyl) phosphate	17	1284.649648	18	1739.150573	21	418.2002481
80	126738	Tributyl phosphate	10	10084.8192	8	9831.148364	17	1132.325665
81	26761400	Diisodecyl phthalate	11	3083.158494	16	2065.731166	8.5	3766.913872
82	26761400 (2)	Diisodecyl phthalate	15	2023.569795	20	1560.110501	8.5	3766.913872
83	27554263	Diisooctyl phthalate	18	1179.157304	17	1918.892147	2	10911.11788
84	28553120	Diisononyl phthalate	14	2770.327291	12	2696.204191	7	5891.650307
88	3648202	Diundecyl phthalate	19	1118.372515	21	638.3397129	22	349.8429931
89	68515515	1,2-Benzenedicarboxylic acid, di-C6-10-alkyl esters	21	996.4809809	15	2139.888363	20	517.1081414
92	84742	Dibutyl phthalate	13	2827.494675	10	6357.753707	5	7248.869072
93	85687	Butyl benzyl phthalate	7	17117.86533	3	25694.52026	3	8911.621274
159	84753	DIHEXYL PHTHALATE	20	1019.036216	19	1720.967781	4	7727.830542

Appendix E – Isomeric structures

The table contains the isomeric structures of the isomers in the substance list represented by SMILES notations.

#	CAS	Name	SMILES
18	32536-52-0	OCTABROMODIPHENYL ETHER	Brc1cc(Br)c(Br)cc1Oc1c(Br)c(Br)c(Br)c(Br)c1Br
19	32536-52-0	OCTABROMODIPHENYL ETHER	Brc1cc(Oc2cc(Br)c(Br)c(Br)c2Br)c(Br)c(Br)c1Br
53	25973-55-1	2-(2H-Benzotriazol-2-yl)-4,6-ditertpentylphenol	CCC(C)(C)c1cc(C(C)(C)CC)cc(N2Nc3ccccc3N2)c1O
95	25973-55-1 (2)	2-(2H-Benzotriazol-2-yl)-4,6-ditertpentylphenol	CCC(C)(C)c1cc(C(C)(C)CC)cc(N2N=C3C=CC=CC3=N2)c1O
59	25637-99-4	Hexabromocyclododecane	BrC1(Br)CCCCCCCCCC(Br)(Br)C1(Br)Br
60	25637-99-4 (2)	Hexabromocyclododecane	BrC1CCC(Br)C(Br)CCC(Br)C(Br)CCC1Br
68	1330-78-5	Tritolyl phosphate	Cc1cccc(O)c1OP(=O)(Oc1ccc(O)cc1C)Oc1cc(O)ccc1C
69	1330-78-5 (2)	Tritolyl phosphate	Cc1ccc(OP(=O)(Oc2ccc(C)cc2)Oc2ccc(C)cc2)cc1
70	1330-78-5 (3)	Tritolyl phosphate	Cc1ccccc1OP(=O)(Oc1ccccc1C)Oc1ccccc1C
72	25155-23-1	TRIXYLYL PHOSPHATE	Cc1cccc(OP(=O)(Oc2ccc(C)cc2C)Oc2cc(C)cc(C)c2)c1C
73	25155-23-1 (2)	TRIXYLYL PHOSPHATE	Cc1cc(C)cc(OP(=O)(Oc2cc(C)cc(C)c2)Oc2cc(C)cc(C)c2)c1
74	26444-49-5	Cresyl diphenyl phosphate	Cc1ccc(OP(=O)(Oc2ccccc2)Oc2ccccc2)cc1
75	26444-49-5 (2)	Cresyl diphenyl phosphate	Cc1ccccc1OP(=O)(Oc1ccccc1)Oc1ccccc1
81	26761-40-0	Diisodecyl phthalate	CC(C)CCCCCCCOC(=O)c1ccccc1C(=O)OCCCCCCCC(C)C
82	26761-40-0 (2)	Diisodecyl phthalate	CC(C)(C)CCCCCCOC(=O)c1ccccc1C(=O)OCCCCCCC(C)(C)C
100	1118-46-3	Butyltin trichloride	CCCC[Sn](Cl)(Cl)Cl
101	1118-46-3 (2)	Butyltin trichloride	CCCC[Sn]{3+}(.Cl{-})(.Cl{-}).Cl{-}
114	683-18-1	Dibutyltin dichloride	CCCC[Sn](Cl)(Cl)CCCC
115	683-18-1 (2)	Dibutyltin dichloride	CCCC[Sn]{2+}(.Cl{-})(.Cl{-})CCCC
116	77-58-7	Dibutyltin dilaurate	CCCCCCCCCCCC(=O)O[Sn](CCCC)(CCCC)OC(=O)CCCCCCCCCCC
117	77-58-7 (2)	Dibutyltin dilaurate	CCCCCCCCCCCC(=O)O{-}.[Sn]{2+}(CCCC)(CCCC).O{-}C(=O)CCCCCCCCCCC
118	78-04-6'	Dibutyltin maleate	CCCC[Sn]1(CCCC)OC(=O)C=CC(=O)O1
119	78-04-6' (2)	Dibutyltin maleate	CCCC[Sn]{2+}1(CCCC).O{-}C(=O)C=CC(=O)O{-}.1