ASSESSMENT AND MODELLING OF THE TOXICITY OF PHENOLS: A COMPARATIVE STUDY WITH MARINE AND FRESHWATER ALGAE

by

Murat Doğa Ertürk

B.S. in Environmental Engineering, Istanbul Technical University, 2003

M.S. in Environmental Sciences, Marmara University, 2005

Submitted to the Institute of Environmental Sciences in partial fulfillment of the requirements for the degree of

Doctor

of

Philosophy

in

Environmental Sciences

Boğaziçi University
2013

ACKNOWLEDGEMENTS

I would like to thank everyone who has, in one way or another, contributed to the completion of this Ph.D. thesis. I'm particularly grateful to my thesis advisor, Prof. Dr. Melek Türker Saçan, for inspirational discussions and her understanding. I'd also like to thank to the jury members, Prof. Dr. Ferhan Çeçen, Prof. Dr. Nilsun İnce, Prof. Dr. Safiye Sağ Erdem and Assist. Prof. Dr. Başak Güven for their valuable comments. I'd like to offer my sincere gratitude to Prof. Dr. Nilsun İnce for kindly allowing me to use the gas chromatograph, to Prof. Dr. Miray Bekbölet for the use of spectrofluorimeter and to Prof. Dr. Işıl Balcıoğlu for the use of several instruments in her lab during this thesis. A very special thanks to Prof. Dr. Andrzej Furman, in particular for the statistics course which helped me immensely throughout the thesis. I'm also thankful to Gülhan Özkösemen for her help during the characterization of seawater samples. Sincere thanks to my friends, Emrah Çoraman, Öncü Maracı, and all my colleagues in the Institute of Environmental Sciences.

I gratefully appreciate the financial supports of Boğaziçi University Scientific Research Funds (Projects 09Y105P and 5564) and TÜBİTAK bilateral project (No: 108Y119). I'm grateful to Prof. Dr. Marjana Novic and Dr. Nikola Minovski for instructing me how to construct counter propagation artificial neural networks.

Lastly, I'd like to thank to all the cronopios I've known so far for sharing their presence with mine and, of course, to Karga for hosting most of these interesting characters. The most special thanks go to the most beautiful cronopio, the one and only, Şule Ertürk. Finally, I'd like to thank my parents, Asuman and Mehmet, for their never-ending love.

ABSTRACT

Algae sustain biodiversity in aquatic ecosystems by producing oxygen and recycling nutrients. In contrast to their key role in the environment, toxicity data of many organic pollutants, such as phenols, on algae, especially for marine algae, are severely limited. On the other hand, the data requirement in algal toxicity is almost impossible to be supplied through exhaustive laboratory testing considering the huge number of chemicals to be assessed. Therefore, the use of alternative methods to laboratory testing, such as the quantitative structure-activity/toxicity relationships (QSARs/QSTRs), can help reduce the data gap in algal ecotoxicity.

In this study, novel toxicity data of phenols on marine alga *Dunaliella tertiolecta* and freshwater alga *Chlorella vulgaris* were generated and subjected to QSAR analysis. The phenols selected for toxicological assessment are known to elicit toxicity through different modes of toxic action including polar narcosis, respiratory uncoupling and reactive mechanisms; as such, the data set was regarded as a miniature model of industrial chemical space and provided a realistic basis upon which to explore the development of algal QSTRs. Multiple linear regression and counter propagation artificial neural network techniques were used to build internally and externally validated QSTR models. Most of the QSTRs highlighted the importance of hydrophobicity and electrophilicity related parameters among numerous descriptors. Hydrophobicity was found to underpin the toxicity of phenols to algae. On the other hand, pyrogallol, hydroquinones and catechols, which are potentially capable of being oxidized to reactive species, displayed algal toxicity in excess of that predicted by hydrophobicity. The toxicity of these reactive phenols was better described by electrophilicity parameters. The external validation of the models was also verified using a data set obtained from literature comprising the toxicity of phenols and anilines to another freshwater alga, *Pseudokirchneriella subcapitata*. Consequently, the developed QSTRs were shown to be applicable to data from another algal test system and at least for another class of organic compounds.

Apart from the QSTRs, investigation of inter-algal and inter-species toxicity correlations between algae and other aquatic organisms such as bacteria, protozoa, daphnia and fish revealed that the response of aquatic organisms to phenols differentiated above the level of polar narcosis. As a result, for a heterogeneous set of compounds acting through different modes of toxic action, the models developed in this study can be used to predict the toxicity of untested compounds provided that the new chemicals are within the applicability domain of the respective model.

ÖZET

Algler gerek oksijen üreterek, gerekse besin maddelerinin çevrimindeki rolleri nedeniyle sucul ekosistemlerde biyolojik çeşitliliğin devamını sağlarlar. Alglerin çevredeki rollerine kıyasla, fenoller gibi birçok organik kirleticilere ait alg toksisite verisi oldukça sınırlıdır. Öte yandan, bütün kimyasallara ait alg toksisite verisini laboratuvar deneyleriyle elde etmek neredeyse imkansızdır. Bu nedenle, kantitatif yapı-aktivite/toksisite ilişkileri (KYAİ/KYTİ) gibi laboratuvar deneylerine alternatif yöntemler alg ekotoksisite verisindeki boşluğun doldurulmasına katkı sağlayabilir.

Bu çalışmada, fenollerin bir tuzlu su alg türü olan *Dunaliella tertiolecta* ve bir tatlı su alg türü olan *Chlorella vulgaris* üzerindeki toksik etkileri belirlenmiş ve elde edilen veriler KYTİ yöntemiyle analiz edilmiştir. Toksisite değerlendirmesi için seçilen fenoller polar narkotik, enerji metabolizmasını hedefleyen ve reaktif fenoller olmak üzere çeşitli mekanizmalarla toksik etki gösterdikleri için endüstriyel kimyasalların minyatür bir modelini temsil etmekte, dolayısıyla da KYTİ analizleri için gerçekçi bir temel teşkil etmektedirler. Çoklu doğrusal regresyon ve tersine yayılmalı yapay sinir ağları teknikleri kullanılarak içsel ve harici olarak geçerliliği test edilen çeşitli KYTİ modelleri elde edilmiştir. Geliştirilen modellerin çoğunda, çok sayıda tanımlayıcı arasında, hidrofobisite ve elektrofilisite bazlı parametrelerin önemi dikkat çekmiştir. Diğer yandan, pirogalol, hidrokinon ve katekol gibi reaktif türlere yükseltgenebilen fenollerin alg toksisiteleri hidrofobisitenin tahmin ettiğinden daha fazla bulunmuştur. Bu reaktif fenollerin toksisiteleri elektrofilisite parametresi ile daha iyi açıklanmıştır. Modellerin harici geçerliliği, fenol ve anilin türevlerinin bir tatlı su alg türü olan *Pseudokirchneriella subcapitata* üzerindeki literatür toksisite verisi ile de doğrulanmıştır. Dolayısıyla, geliştirilen modellerin farklı bir alg türüne ve en azından farklı bir organik kirletici grubuna da uygulanabilir olduğu gösterilmiştir.

KYTİ modellerinin yanı sıra, algler ile bakteri, protozoa, su piresi ve balık gibi sucul türler arasında yapılan toksisite korelasyonlarından çıkan sonuç, sucul organizmaların

fenolere verdiği tepkilerin polar narkotik mekanizmanın üzerinde farklılaşmaya başladığıdır. Sonuç olarak, bu çalışmada geliştirilen KYTİ modelleri, modellerin uygulanabilirlik alanında kalması şartıyla, farklı toksik mekanizmalar uyarınca etki gösteren ve deneysel verisi olmayan çeşitli kimyasalların toksisitesini tahmin etmek için kullanılabilir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

Symbol	Explanation	Unit
CP	Chlorophenol	
DCP	Dichlorophenol	
Cat	Catechol	
Cv	*Chlorella vulgaris*	
cv	Cross validation	
DMSO	Dimethyl sulfoxide	
Dm	*Daphnia magna*	
Dt	*Dunaliella tertiolecta*	
EC_{50}	Effective concentration that reduces the observed endpoint by 50%	mg/L
Hq	Hydroquinone	
IC_{50}	Concentration that inhibits algal growth by 50%	mg/L
ICp	Linear interpolation combined with bootstrapping	
LOEC	Lowest observed-effective concentration	mg/L
Log D	Distribution coefficient	
Log Kow	Logarithm of *n*-octanol/water partition coefficient	
Log P	Partition coefficient	
mM	Milimolar	
MOA	Mode of action	
NOEC	No observed-effect concentration	mg/L
PCP	Pentachlorophenol	
Pp	*Pimephales promelas*	

P.R.	Polynomial regression	
Ps	*Pseudokirchneriella subcapitata*	
*p*T	Negative logarithm of EC_{50}/IC_{50}	mmol/L
QSAR	Quantitative Structure-Activity Relationship	
QSTR	Quantitative Structure-Toxicity Relationship	
Res	Resorcinol	
SGR	Specific growth rate	d^{-1}
TeClHq	Tetrachlorohydroquinone	
TCP	Trichlorophenol	
TeCP	Tetrachlorophenol	
Tp	*Tetrahymena pyriformis*	
Vf	*Vibrio fischeri*	

1. INTRODUCTION

Information on aquatic toxicity is required for various trophic levels to properly evaluate the possible risks that chemicals may pose to the environment. As such, evaluation of data from algal growth inhibition tests has been an integral part of environmental risk assessment (Christensen et al., 2009) which is inherent in the new chemical management system called REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) adopted by the European Council on 18 December 2006. According to REACH, all substances on the European market, which are manufactured or imported in quantities greater than 1 tonne per annum (tpa), have to be evaluated for their adverse effects on the environment. In the evaluation process for those chemicals, REACH requires basic ecotoxicological information including short-term toxicity data on green algae (EC, 2006).

Considering the vast number of chemicals that needs to be evaluated, REACH promotes the use of alternative methods to laboratory testing for chemical risk assessment (EC, 2006). To this end, computational models, in particular the quantitative structure–activity relationships (QSARs), are expected to be used in filling the data gaps (Netzeva et al., 2008). QSAR as a methodology is based on the premise that the biological activity (e.g., toxicity) of a given compound can be predicted from its molecular structure. Since the structural properties of compounds are usually encoded in a form of molecular descriptors, QSAR can also be defined as a mathematical relationship between the molecular descriptors and the activity of compounds. Consequently, the results obtained are presented in the form of mathematical model(s) which can be used to predict the toxicity of compounds whose toxicity is not known (Saçan et al., 2013).

From an ecological perspective, algae form the base of food webs, they provide food for higher trophic levels and produce oxygen which is vital for the sustainability of life in aquatic environments. Apart from the crucial role they play in the ecosystems, their ubiquitous distribution throughout the globe, ease of collection and culturing, and rapid growth rate make them ideal for laboratory testing (DeLorenzo, 2009). Despite their vital

functions in the aquatic ecosystems and advantages for laboratory testing, reliable algal toxicity data are limited (Cronin et al. 2004; Netzeva et al., 2008). As recently demonstrated by Aruoja et al. (2011), the number of algal QSARs available in the literature are remarkably low compared to those constructed using other aquatic species such as fish, protozoa, bacteria or daphnia. This is probably a reflection of the lack of reliable algal data in the literature (Cronin et al., 2004).

The toxicity data on marine algae is even scarcer. The reasons behind the lack of marine ecotoxicity data in general has been attributed to the fact that urbanization has rendered freshwater ecosystems more vulnerable to the chemical releases which in turn made the protection of freshwater communities the first priority by regulatory schemes. Consequently, there is substantial amount of information available on the toxicity of chemicals to freshwater organisms while there are relatively fewer data on the effect of chemicals to marine organisms, in particular to aquatic plants and algae (ECETOC, 2001; ECETOC, 2008). Once chemicals are released into the environment, they can cause adverse effects in non-target organisms in marine environments. Therefore, determination of the toxicity of chemicals to non-target species such as marine algae, especially in countries like Turkey which is surrounded by sea from three sides, will be beneficial to understand the impact of chemicals to marine ecosystems (Ertürk and Saçan, 2012). In this thesis, phenols were chosen for toxicological assessment considering their widespread use and environmental significance as phenols have been detected in terrestrial and aquatic food chains (Jensen, 1996) and in environmental samples, particularly in those obtained from aquatic ecosystems (WHO, 1987; WHO, 1989; WHO, 1994). The first aim of this thesis was, therefore, to determine the toxicity of selected phenols to marine alga, *Dunaliella tertiolecta*. Taking into account the vast number of chemicals that needs to be evaluated for their potential adverse effects, the data gap in marine algal ecotoxicity is almost impossible to fill with exhaustive testing of each chemical. As the experimental determination of toxicity is a costly and time consuming process, QSARs can be used to predict the toxicities of untested chemicals. Therefore, the second aim of this thesis was to build marine algal QSARs using the novel marine algal toxicity data generated in this study. The developed marine QSARs can be used to fill the data gap present in marine algal toxicity data since they will be fully validated according to the OECD criteria.

The traditional use of QSAR in ecotoxicology and environmental risk assessment is to predict the toxicity of untested chemicals. However, QSAR methodology can also be used as a tool to predict the toxicity of a compound towards a particular biological organism using toxicity data on another species (Zvinavashe et al., 2008). In fact, possibilities for these extrapolations have been demonstrated previously in the literature for industrial chemicals (Kaiser, 1998; Dimitrov et al., 2000), even for pharmaceuticals (Kar and Roy, 2010). However, not much attention has been directed to algae. Therefore, to be able to compare the responses of freshwater and marine algae, the third aim of this thesis was to determine the toxicity of selected phenols to freshwater alga, *Chlorella vulgaris* and the fourth aim was to develop algal QSARs using the toxicity data generated for *C. vulgaris.* Finally, as the fifth and final aim, was to predict the toxicity of phenols to *D. tertiolecta* using toxicity of the same phenols on freshwater alga, *C. vulgaris*. The final aim also covers a risk assessment perspective by evaluating the interspecies toxicity correlations in an attempt to fill the data gaps in algal ecotoxicity. In addition, the evaluation of the predictive and classification performance of the widely used ECOSAR model was also attempted.

2. THEORETICAL BACKGROUND

Aquatic ecosystems have been severely threatened by intentional and accidental discharges of toxic compounds. From a historical perspective, the counter measures to prevent aquatic pollution due to the adverse effects of chemicals resulted in the acceptance of global physico-chemical and bio-chemical environmental parameters. However, as the experience in the field of risk assessment has increased, it was perceived that parameters such as the chemical oxygen demand or the biological oxygen demand were not adequate to provide necessary information on the potential harmful effects of chemicals to the aquatic environment (Saçan and Balcıoğlu, 2006). Since living organisms respond quickly to disruptions in their habitats, biological assays have become important tools in assessing the environmental impact of chemicals and thereby assisting in the development of precautionary measures and strategies for environmental management (Sponza, 2002). However, ever-increasing chemical usage for domestic, agricultural and industrial purposes has posed another problem to be resolved by the scientific community and decision-makers: the need to develop fast and reliable methods to evaluate the potential threat posed by chemicals on the ecosystems.

On 18 December 2006, the European Council and European Parliament adopted a new legislation for a chemical management system called REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) (EC, 2006). According to REACH, all substances on the European market, which are manufactured or imported in quantities greater than 1 tonne per annum (tpa), have to be evaluated for their adverse effects on the environment by the year 2018. Based on a recent evaluation, the number of such chemicals is between 68000 and 101000[1] (Costanza and Hartung, 2009). According to REACH, the basic ecotoxicological information for those substances (i.e., imported or manufactured > 1 tpa) include short-term toxicity testing on green algae. Since toxicity testing of each chemical -either in the market or in the pipeline- seems to be an almost impossible task to fulfill in such a limited time, the legislation also includes the promotion of alternative methods to laboratory testing for chemical risk assessment (EC, 2006).

[1] The number of substances to be evaluated changes with time. The estimated figures can be followed from the European Chemicals Agency (ECHA) website (http://echa.europa.eu/).

Computational models for toxicity prediction are among the alternative methods to laboratory testing. In this context, these rapid and low cost methods are attractive to use in filling data gaps and for classification and labeling purposes (Aptula et al., 2005a). In particular, structure – activity relationships (SARs) and quantitative structure–activity relationships (QSARs), collectively referred to as (Q)SARs, are expected to be useful for substances lacking in experimental data (Netzeva et al., 2008).

2.1. Quantitative Structure – Activity Relationships (QSARs)

The quantitative relationship between molecular structure and the biological activity (e.g., toxicity) is the center of focus for the field of quantitative structure–activity relationships (QSARs). For any given QSAR, the main objective is to construct and explore these quantitative relationships through mathematical models (Roy et al., 2008).

The QSAR history dates back to the earliest report on a quantitative relationship between molecular and biological properties reported by A. F. A. Cros in 1863, where the researcher observed an increase in the toxicity of alcohols to mammals with decreasing water solubility (Cros, 1863 as cited in Kubinyi, 2002). In the years that followed, one of the most important contributions in the early days of QSAR came from L.P. Hammett around 1935 when the researcher postulated that the effect of substituents on the ionization of benzoic acids could be used as a model system to estimate the electronic effect of substituents on similar reaction systems (Hansch and Leo, 1995). Despite further contributions from several researchers in those early days, around 1960, QSAR methodology seemed to arrive to a dead end. However, in 1964, C. Hansch, T. Fujita, S. M. Free, Jr. and J. W. Wilson moved the field forward and started the classical QSAR which is still in use today (Kubinyi, 2002). Among these researchers, the contributions of Corwin Hansch through his investigations on the relationship between lipophilicity and biological activity merits further introduction.

In 1962, Hansch and co-workers published a brilliant study on the structure-activity relationships of plant growth regulators and their dependency on Hammett constants and hydrophobicity (Hansch et al., 1962 as cited in Selassie, 2003). Using the octanol/water

system, a whole series of partition coefficients was measured, and thus a new hydrophobic scale was introduced (Kubinyi, 2002; Selassie, 2003):

$$\pi_x = \log P_x - \log P_H \quad (2.1)$$

where π_x denotes the lipophilicity contribution of a substituent x, based on hydrogen as a reference substituent. Hansch and Fujita (1964) then combined the hydrophobic constants (π) with Hammett's electronic constants (σ) to yield the so-called Hansch equation (also called linear free energy relationship) and its many extended forms (Kubinyi, 2002 and references therein):

$$\text{Log } 1/C = a\sigma + b\pi + \text{constant} \quad (2.2)$$

where C is a molar dose that produces a certain biological response, a and b are the coefficients for σ and π, respectively.

The development of these models led to an unprecedented increase in the number of QSAR studies and related approaches such as the extensive studies on molecular connectivity by Kier and Hall (1976). The milestones in the early development of QSAR methodology as well as other methods that have been developed since Hansch analysis can be found in the reviews on QSAR history by Kubinyi (2002) or in the study by Selassie (2003). In summary, a great deal of work in the area of QSAR has amassed in the literature since the contributions of Hansch and Fujita (1964).

As noted earlier, the basic premise of a QSAR is that the activity of a compound can be predicted by its structural/physico-chemical characteristics. In toxicological research, the term activity is generally used interchangeably with toxicity; therefore, the acronym QSAR is often renamed as QSTR (quantitative structure-toxicity relationship). Consequently, the models developed in this thesis are more correctly defined as QSTRs. In a QSTR study, the relationship between structure and toxicity is studied through theoretical or experimental values called descriptors, which simply capture each compound's structural, physico-chemical, or quantum chemical property. Basically, any given QSTR can be divided into three basic components:

a) the biological toxicity data (i.e., the dependent variable),
b) the descriptor(s) (i.e., independent variable(s)), and
c) the mathematical, or statistical, relationship between the dependent and independent variables.

2.1.1. The biological (toxicity) data

Information on aquatic toxicity is required in order to assess the risk of chemical substances to marine and freshwater organisms living in the water column. An adverse effect of a chemical can be evaluated after short term and/or long-term exposure. The short-term (acute) toxicity is usually measured as the concentration that decreases a measurable effect (e.g., growth rate) by 50% (Effective Concentration, EC_{50}; Inhibitory Concentration, IC_{50}) (Netzeva et al, 2008). Prior to a QSTR analysis, the toxicity data in the form of $LC_{50}/EC_{50}/IC_{50}$ (in mmol/L) is usually subjected to a negative log transformation in order to obtain normal distribution of data. By this means, the higher numerical value corresponds to higher toxicity (which is the opposite when EC_{50} values are considered, where a compound with lower EC_{50} value is more toxic than the one with higher EC_{50}).

In this thesis, the toxicity experiments have been carried out according to standardized test protocols (OECD, 2006; APHA-AWWA-WEF, 1998). Algae were chosen as the biological organisms to be used in ecotoxicity testing for several reasons. First of all, from an ecological perspective, algae form the base of food webs (i.e., primary producers), they provide food for higher trophic levels and produce oxygen which is vital for the sustainability of life in aquatic environments (Saçan et al., 2013). The importance of algae is better understood considering the fact that approximately half of the planetary primary production is generated by marine phytoplankton, which also affects the abundance and diversity of marine organisms. Furthermore, algae strongly influence biochemical cycles, such as nitrogen and carbon cycles (Boyce et al., 2010). Apart from the crucial role they play in the aquatic ecosystems, their ubiquitous distribution throughout the globe, ease of collection and culturing, high surface area to volume ratio and rapid growth rate make them ideal for laboratory testing (De Lorenzo, 2009). Lastly, as indicated by Cronin et al. (2004), while there are relatively large toxicity databases for the fish and crustaceans,

which represent higher trophic levels, limited toxicity data are available for the primary producers. In a recent bibliographic investigation carried out by Aruoja et al. (2011), it was demonstrated that most of the QSTR models have been constructed using toxicity data on fish and the protozoan *Tetrahymena*, followed by *Daphnia* and the bacterium *V. fischeri*. As stated by the researchers, the QSARs/QSTRs that have been developed using algal data are remarkably low (Aruoja et al., 2011). On the other hand, REACH requires ecotoxicological information for substances manufactured or imported in quantities of 1–10 tonnes per year. The basic ecotoxicological information required within REACH includes the acute toxicity testing on algae (EC, 2006). However, algal toxicity data compatible with the requirements of REACH can only be found in a few datasets (Aruoja et al., 2011).

The biological data to be modelled should be of high quality to develop a quality QSAR/QSTR. In other words, the accuracy of any QSAR/QSTR model depends on the quality of data that has been used for its development (Selassie, 2003). In this thesis, the toxicity experiments have been conducted in the same laboratory, by the same researcher, following the standardized test protocols; therefore, the aquatic toxicity data generated in this thesis should meet the major requirements for high quality data. Therefore, the QSTR models developed in this thesis are expected to be used in a risk assessment scenario for green algae.

On the basis of ecological and practical reasons as well as taking regulatory needs into consideration, this thesis was structured to generate new toxicity data using freshwater and marine algae. Several *in silico* strategies, most importantly QSTRs and interspecies toxicity correlations, were developed and evaluated to help reduce the data gap present in algal ecotoxicity.

2.1.1.1. Toxicity testing with algae. Evaluation of data from toxicity tests using microalgae is an integral part of environmental risk assessment (Christensen et al., 2009). The algal toxicity data are required because algae are a key functional group of organisms. From an ecological perspective, the toxic stress may alter the composition of the phytoplankton community which in turn might affect the structure and functioning of the whole ecosystem. Additionally, the possibility that specific toxic effects may be better expressed

in algae in low concentrations makes this test indispensible from a risk assessment perspective (Nyholm and Källqvist, 1989).

The purpose of toxicity testing with algae is to determine the effects of a chemical on the growth of these organisms (OECD, 2006). The response obtained from algal tests is a continuous parameter not influenced by the individual tolerance of the test organisms, because cells from a single clone are applied in great number (Christensen et al., 2009). The basic idea behind the tests is to expose exponentially growing test organisms to the test substance in batch cultures over a prescribed test period (usually 48 to 96 hours). Despite brief test durations, effects over several generations can be assessed. The system response is the reduction of growth in a series of algal cultures exposed to increasing concentration of a test substance where the algal growth is quantified from measurements of algal biomass as a function of time. Due to the difficulties in determining the dry weight per volume of the algal biomass, surrogate parameters are used which include cell counts, optical density, fluorescence, etc., to quantify algal biomass. The test endpoint is the inhibition of growth, expressed as the logarithmic increase in biomass (average specific growth rate) during the exposure period. From the average specific growth rates recorded in a series of test solutions, the concentration bringing about a specified x % inhibition of growth rate (e.g., 50%) is determined (OECD, 2006). An additional response variable used in the most recent guideline prepared by the OECD (2006) is yield, which is defined as the biomass at the end of the exposure period minus the biomass at the start of the exposure period. From the yield recorded in a series of test solutions, the concentration bringing about a specified x % inhibition of yield (e.g., 50 %) is calculated. Based on the data generated in algal tests, the chemicals are ranked with regard to their environmental toxicity for use in environmental hazard evaluations (Nyholm and Källqvist, 1989).

For the toxicity tests to be carried out in this thesis, two unicellular algae, one representative of marine (*Dunaliella tertiolecta*) and the other one representative of freshwater (*Chlorella vulgaris*) environments, were selected.

2.1.1.2. *Dunaliella tertiolecta*. Unicellular green algae of the genus *Dunaliella* have been studied as a model organism since the early 19th century. *Dunaliella* species are ubiquitous in saline environments and their wide distribution may be a result of their tolerance to a

wide range of salinities, light intensities, temperatures as well as the ability to survive for years among salt crystals (Polle et al., 2009). Lack of a rigid cell wall in *Dunaliella* has been regarded as an excellent opportunity to study the growth response of algae to environmental conditions and toxicants (Rao, 2009).

Dunaliella tertiolecta (Figure 2.1) is a motile species and, similar to other species in the genus, has a high tolerance for salt, temperature, and light. Since *D. tertiolecta* cells do not clump together or form chains, the morphology makes it ideal to perform cell counts (NIWA, 1998). It has been regarded as a species representative of marine environments and has been recommended as a model organism for ecotoxicological studies (APHA, AWWA, WEF, 1998; DeLorenzo, 2009). The scientific classification of *Dunaliella tertiolecta* according to González et al. (2009) is provided in Table 2.1.

Table 2.1. Scientific classification of *Dunaliella tertiolecta*

Domain	Eukaryota
Kingdom	Viridiplantae
Division	Chlorophyta
Class	Chlorophyceae
Order	Chlamydomonadales
Family	Dunaliellaceae
Genus	Dunaliella
Species	*Dunaliella tertiolecta*

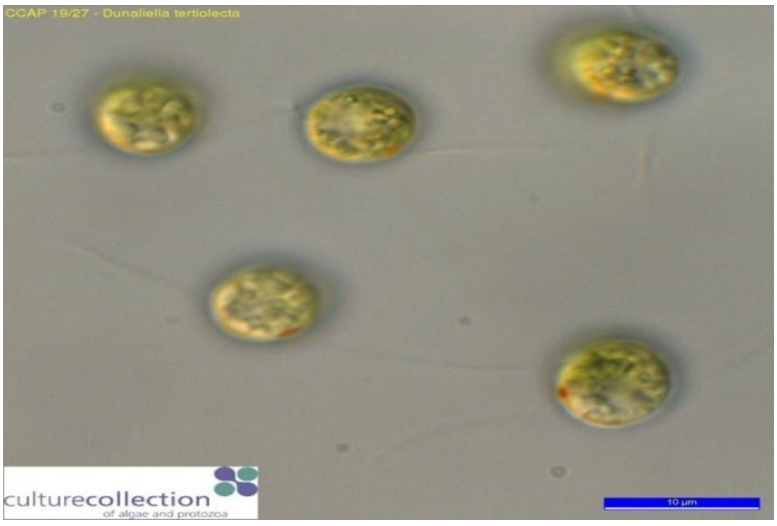


Figure 2.1. Microscopic view of *Dunaliella tertiolecta*
(© Culture Collection of Algae and Protozoa – with permission)

Despite its suitability as a model test species, the number of studies that aimed to determine the toxicity of organic compounds towards *D. tertiolecta* is limited. Consequently, this resulted in the lack of an available QSAR/QSTR based on toxicity data using *D. tertiolecta* as the test species.

A recent query (only for compounds with a reported growth inhibition endpoint) in the ECOTOX database (last access: December 2012) revealed several important studies aimed to determine the toxicity of organics towards *D. tertiolecta*. Among the studies that were obtained from the search in the ECOTOX database, the following studies were highlighted and deemed particularly important because the bioassays therein were conducted in a single laboratory by a standard protocol. Most of these studies were conducted by DeLorenzo and co-workers where the researchers has determined the toxicity of various pesticides and personal care products to *D. tertiolecta* (DeLorenzo and Serrano, 2006; DeLorenzo and Fleming, 2008). Additionally, Wiegman et al. (2001) determined the toxicity of azaarenes, a family of N-heterocyclic polycyclic aromatic hydrocarbons using *D. tertiolecta*. Okumura et al. (2001) investigated the toxicity of several organic solvents to marine algae, including *D. tertiolecta*. The only compound used in this thesis having a literature toxicity data on marine alga *D. tertiolecta* is pentachlorophenol with a 96 h IC_{50} of 0.17 mg/L (ECOTOX, 2012).

2.1.1.3. *Chlorella vulgaris*. The genus *Chlorella* is comprised of green algae which are non-motile, globular and unicellular with an average diameter of 4-10μm (Kuhl and Lorenzen, 1964). Small, spherical or elliptical coccoid green algae without any special morphological features such as spines or bristles have mostly been assigned to the genus *Chlorella* Beijerinck. Since his description of *Chlorella vulgaris* (Beijerinck, 1890, as cited in Darienko et al. 2010) more than a hundred species have been established. The scientific classification of *C. vulgaris* is provided in Table 2.2.

Table 2.2. Scientific classification of *Chlorella vulgaris*

Domain	Eukaryota
Kingdom	Plantae
Division	Chlorophyta
Class	Trebouxiophyceae
Order	Chlorellales
Family	Chlorellaceae
Genus	Chlorella
Species	*Chlorella vulgaris*

Figure 2.2. Microscopic view of *Chlorella vulgaris*
(© Culture Collection of Algae and Protozoa – with permission)

Chlorella vulgaris has been selected as a test organism in several toxicity studies (Scragg, 2006; Cai et al., 2009; Sahinkaya and Dilek, 2009; Murkovski and Skórska, 2010), because of its widespread distribution and natural presence in freshwater ecosystems (Ventura et al., 2010). However, in contrast to its abundance, the toxicity data for *C. vulgaris* has significant gaps. In an effort to fill this data gap, Worgan et al. (2003) proposed a fast screening method to determine the toxicity of organic compounds to *C. vulgaris*. This method is a 15-minute assay, based on the premise that all organisms, including algae, contain non-specific esterases, the activity of which can be assessed by the measurement of the disappearance of an ester, or the appearance of the product. Using this methodology, Cronin et al. (2004) developed a toxicity dataset of 91 organic compounds to *C. vulgaris*. To the best of our knowledge, this is the largest toxicity dataset generated in a single laboratory using *C. vulgaris*. Although the premise of the test allows fast screening

of toxicants, it should be noted that algae are known to adapt to low concentrations of toxicants; therefore, as also suggested by the developers of the test design (Worgan et al., 2003), its applicability should be verified with data obtained from standardized algal growth inhibition assays. In an earlier study, Kramer and Trümper (1986) reported the *in vitro* toxicological analysis of a series of mono-substituted phenols to *C. vulgaris*. The test was performed using six hour of exposure under static conditions. Similar to the 15-min assay as described by Worgan et al. (2003), the applicability of data produced by Kramer and Trümper (1986) should also be verified against data obtained from standardized algal growth inhibition assays. Apart from these two assays utilizing endpoints other than growth, the toxicity of phenol and eight chlorophenols towards *C. vulgaris* was previously determined by Shigeoka (1988 - as retrieved from TerraTox™, 2006) in algal growth inhibition assays. It should be noted that, although the toxicity of eight chlorophenols towards *C. vulgaris* was determined previously, the toxicity of all chlorophenols as well as polyphenols such as pyrogallol, hydroquinones, catechols and resorcinols were determined for the first time in this thesis. Consequently, the phenolic data set presented in this thesis enables the development of comprehensive freshwater algal QSARs/QSTRs.

2.1.2. The descriptors

Molecular descriptors are simply the mathematical representations of a molecule obtained by a strict and well-defined algorithm which is applied to a defined molecular representation or a specified experimental procedure (Consonni and Todeschini, 2010). In other words, a molecular descriptor can be seen as "the final result of logic and mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number or the result of some standardized experiment" (Todeschini and Consonni, 2000).

The molecular descriptors can be broadly classified into two major groups: experimental and theoretical descriptors. While the experimental descriptors, as the name suggests, are the numerical data obtained at the end of experimental procedures (Karelson, 2000); the theoretical descriptors are derived from a symbolic representation of the molecule (Consonni and Todeschini, 2010). The experimental descriptors include the experimental measurements such as the logarithm of the *n*-octanol/water partition

coefficient (Log K_{ow} or Log P), molar refractivity, dipole moment, polarizability (Consonni and Todeschini, 2010) as well as other experimental data obtained at the end of thermodynamic, kinetic, spectroscopic measurements (Karelson, 2000). For theoretical descriptors, a general classification is provided in Table 2.3.

The fundamental difference between the experimental and theoretical descriptors is that the former contains error due to experimental noise while the latter does not (Consonni and Todeschini, 2010). Furthermore, depending on the characteristic of the compound, the determination of the experimental descriptor may be difficult, expensive or even impossible for some chemicals (Karelson, 2000). With respect to experimental descriptors, the greatest advantages of theoretical descriptors are that they are easily available (usually in a commercial software), and the procedures to obtain these descriptors are relatively faster and cheaper (Consonni and Todeschini, 2010). This, of course, does not mean that experimental descriptors are not important. As discussed previously, the foundations of the QSAR methodology is based on experimental descriptors, such as the partition coefficient, Log P, that has been used as a measure of hydrophobicity for over a century (Cronin and Schultz, 2003). In this thesis, both theoretical and experimental descriptors were used to construct QSTRs provided that they offer a mechanistic interpretation for the developed relationships.

The descriptor list provided in Table 2.3 does not include all the available descriptor classes. As stated by Consonni and Todeschini (2010), the number of descriptors is continuously increasing in direct proportion with the increasing interest for deeper investigations on chemical and biological systems. Detailed information on descriptors can be found in the works of Todeschini and/or Consonni (2000, 2010, 2011).

Table 2.3. General classification of molecular descriptors (from Karelson, 2000)

Class	Subclass
Constitutional descriptors	Counts of atoms or bonds
	Atomic weight-based descriptors
Topological descriptors	Topological (connectivity) indices
	Information-theoretical descriptors
	Topochemical descriptors
Geometrical descriptors	Distance-related descriptors
	Surface area-related descriptors
	Volume-related descriptors
	Molecular steric field descriptors
Charge-distribution-related descriptors	Atomic partial charges
	Molecular electrical moments
	Molecular polarizabilities
	Molecular electrostatic field descriptors
Molecular orbital-related descriptors	Frontier molecular orbital energies
	Bond orders
	Fukui's reactivity indices
Temperature-dependent indices	Thermodynamic functions
	Boltzmann factor-weighted descriptors
Solvational descriptors	Electrostatic energy of salvation
	Dispersion energy of salvation
	Free energy of cavity formation
	Hydrogen bonding descriptors
	Entropy of salvation
	Theoretical linear salvation energy descriptors
Mixed descriptors	Topographical descriptors
	Electrotopological descriptors
	Charged partial area descriptors

2.1.3. Linking toxicity to structure through mathematical models

An important task in science and technology is modeling a property by one or several variables (Varmuza and Filzmoser, 2009). The QSAR paradigm is no exception. Several QSAR modeling algorithms have been introduced which can be broadly classified into two major categories: linear and nonlinear modeling algorithms. In this thesis, the models were constructed using multiple linear regression and counter-propagation artificial neural networks.

2.1.3.1. Multiple linear regression (MLR). Linear regression relates a dependent variable, y, to a single independent variable, x, in a linear manner. The most common method used in linear regression analysis is the least squares regression which aims to find the best fit that minimizes the residual sum of squares (Hanrahan, 2009). Regression with one variable is a special case of regression with many x-variables. With multivariate data, which is usually the case in QSAR studies, the dependent variable can be related to a linear combination of x-variables with the general form:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_n x_n \quad (2.3)$$

where the α values indicate the coefficients which measure the weight of each variable in the model (Varmuza and Filzmoser, 2009). In QSAR studies, this technique is especially powerful provided that certain requirements are met. First, the activity (e.g., toxicity) value has to be divisible into additive terms, each of which corresponds to a single molecular descriptor (Karelson, 2000). Note that this property of MLR was successfully exploited by Hansch and Fujita (1964) when they combined the hydrophobic constants (π) with Hammett's electronic constants (σ) to yield the so-called linear free energy relationships or Hansch analysis. Secondly, the independent variables should depend linearly on the dependent variable (Karelson, 2000). In the cases where such a relationship is not linear in nature, it is still possible to find a functional relationship either using a non-linear transformation of data or using a non-linear technique such as artificial neural networks.

Due to its ease of interpretation and transparency (Cronin and Schultz, 2003), MLR has been the major technique used for constructing QSARs (Selassie, 2003). However, it should be noted that MLR is known to suffer from several drawbacks including over-fitting of data, its dimensionality and inability to work on ill-conditioned data. In situations where the number of descriptors exceeds the number of compounds, a variable reduction prior to modeling is required (Hanrahan, 2009). In this thesis, variable reduction was carried out using the heuristic method embedded in the CODESSA software (version 2.2, 1996, ©University of Florida, Gainesville, USA). Heuristic method has been successfully employed in previous QSAR studies in a variety of topics including the classification of compounds according to their toxic modes of action (Russom et al., 1997), toxicity prediction of chemicals to aquatic species (Eroğlu et al., 2007; Moosus and Maran, 2011)

and modeling skin irritation (Golla et al., 2009). Therefore, it is believed that heuristic search is a reliable method to eliminate redundant information inflicted by insignificant descriptors.

2.1.3.2. Artificial neural networks (ANNs). An artificial neural network (ANN) is inspired by the architecture of biological neurons such as the human brain (Zou et al., 2008). The ANN is composed of a number of interconnected elements called artificial neurons which can be defined as the mathematical/statistical tools constructed to mimic the function of the biological neurons in the brain (Saçan et al., 2013). In other words, an artificial neuron is designed to mimic the function of the neural cells of a living organism. The human nervous system consists of approximately 10^{10} neural cells commonly known as neurons. A simplified picture of a neuron can be seen in Figure 2.3. A typical neuron of a motor complex is composed of a cell body (soma) with a nucleus. The cell body has two types of extensions: the dendrites and the axon. The dendrites receive signals and send those signals to the soma while the axon transmits signals to other neurons. In all living organisms, the signals are very similar, and this similarity suggests that the functioning of brain is not so much dependent on the role of a single neuron, but on the entire ensemble of neurons (Zupan and Gasteiger, 1999). Therefore, considering the analogy between biological and artificial neuron, similar to the biological neurons which are capable of accepting a huge number of information in the form of electrical impulses and to process them in an appropriate way, the neuron must be capable of accepting a huge number of input variables of different nature from the neighboring neurons and to process them in a non-linear way in order to give a result (Saçan et al., 2013).

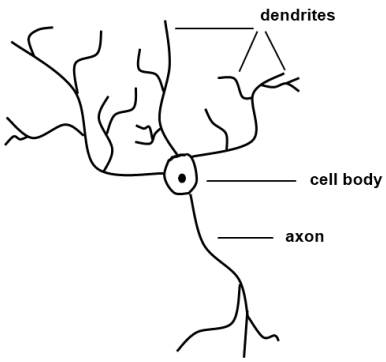


Figure 2.3. Simplified scheme of a neuron

Compared to the linear QSAR modeling strategies, where the relationship between the structure and activity is established in a linear manner, the non-linear QSAR methods such as the ANNs transform the input variables (i.e., descriptors) into corresponding output values (i.e., activity/toxicity) utilizing non-linear transforming functions. Since ANNs are of different types and it is difficult to describe them with only a single definition, they can most easily be presented as a statistical tool, Figure 2.4, consisting of multivariate input $X = (x_1, x_2, x_3,\ldots,x_m)$ and multi-response output $Y = (y_1, y_2, y_3,\ldots,y_n)$ (Saçan et al., 2013).

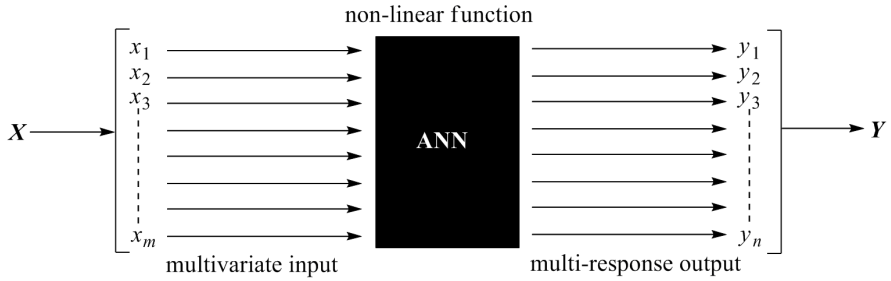


Figure 2.4. Schematic representation of ANNs (taken from Saçan et al., 2013)

ANNs can be constructed using different number of neurons, which form the so-called network architecture. In QSAR modeling, the number of neurons in the ANNs mainly depends on the number of training set objects, which are used to train the network (Saçan et al., 2013). The ANN uses a learning process to train the network. During the learning process the weights are adjusted to desired values. There are two main types of learning approaches: a supervised learning approach and an unsupervised learning approach. In supervised learning, a training set is provided and the weights of neurons are adjusted to minimize the error between the network output and the corrected output. The unsupervised learning procedure, on the other hand, does not use target output values from a training set. The network attempts to uncover the underlying patterns in the data alone (Zou et al., 2008). Supervised learning is generally conducted for data estimation purposes (e.g., prediction) based on previous knowledge on the estimated property, while unsupervised learning is usually used when deriving information from data (e.g., ordination and classification) without previous knowledge. In the latter, solutions are obtained through an adaptive process to reach a global maxima and minima of energy (i.e., information) in the system (Chon, 2011).

The first ANN model in the domain of QSARs was reported almost 40 years ago when Hiller et al. (1971) reported a study dealing with the use of perceptrons, the only type of ANNs known at that time (Hiller et al., 1971, as cited in Baskin et al., 2008). Since then, due to the non-linear nature of biology (Cronin and Schultz, 2003), the ANNs have been frequently used in QSAR studies within the domain of environmental toxicology (Devillers, 2008).

Within the scope of this thesis only one type of ANN, namely the Counter Propagation Artificial Neural Network (CP ANN), was used for the prediction of toxicity of phenols towards algae. However, to better understand the idea behind CP ANNs, a brief introduction to the Kohonen Artificial Neural Networks (KANNs) will be useful. The KANNs or Self Organizing Maps (SOM) represents a basic type of ANNs where the input variables (e.g., descriptors) that form the multi-dimensional space are mapped onto a two-dimensional network of neurons. This mapping procedure progresses through a non-linear algorithm commonly known as training (or learning of the network, as mentioned above). During the learning phase, KANN suits itself to the input objects, so the similar objects are associated with topologically close neurons on the network (Saçan et al., 2013). The nearest neighbor method is one of the most frequently used learning approaches for ANNs (Zou et al., 2008). The Kohonen learning can be literally described as "*winner takes all strategy*". For each input only the most excited (or so-called "*winning neuron*") is selected and therefore, the correction is performed around this "*winning neuron*". The selection of the winning neuron is based on the Euclidean distances between the neurons (between the vectors of weights and objects), i.e., the minimal determined distance gives the winning neuron (Eq. 2.4) (Zupan and Gasteiger, 1999):

$$d_{j,s}^{Eucl} = \sqrt{\sum_{i=1}^{m} \left(W_{j,i} - X_{s,i}\right)^2}, \min\left\{d_{j,s}^{Eucl},\ j = 1 \cdots N_{net}\right\} \rightarrow W_c \qquad (2.4)$$

where W_j are the vectors of weights (w_{j1}, w_{j2}, w_{j3},…, w_{ji},…, w_{jm}), X_s are the objects (x_{s1}, x_{s2}, x_{s3},…, x_{si},…, x_{sm}), N_{net} is the total number of neurons in the network, whereas W_c designates the selected winning neuron. Once the winning neuron is found and selected, the correction of its weights as well as the weights in the neighboring neurons is performed (Eq. 2.5) (Zupan and Gasteiger, 1999):

$$\Delta w_{j,i} = \eta\left(1 - \frac{d}{p+1}\right)\left(x_{s,i} - w_{j,i}^{old}\right), d = 0,1,2,\ldots,p \tag{2.5}$$

where η designates the learning rate, d is topological distance counting rings of neurons between the central neuron (W_c) and the j-th neuron for which the correction should be performed, while p designates the maximum topological distance to which the correction is applied (Eq. 2.6) (Zupan and Gasteiger, 1999):

$$p = \left(1 - \frac{n_{epochs}}{n_{tot}}\right) N_{net} \tag{2.6}$$

At the beginning of the training procedure the correction is applied to the whole network, $n_{epochs} = 0$, and therefore $p = N_{net}$, whereas at the end of the training, $n_{epochs} = n_{tot}$, so $p = 0$, what means that in the last epoch of the learning process, only the weights of the central neuron are corrected. The result of the training represents a two-dimensional array of neurons (also known as top-map), which are occupied with objects in such a way that similar objects are positioned close to each other or on the same neuron (Saçan et al., 2013).

The Counter-Propagation Artificial Neural Networks (CP ANN) can be defined as an extension of KANN (Zupan and Gasteiger, 1999). In comparison with KANN, here the target values (e.g., toxicity) are additionally included in the modeling procedure. Pictorially, they can be represented exactly as KANNs (which contain only one layer of neurons), but extended by another layer of neurons positioned exactly below the Kohonen layer. This additional layer is functioning exactly as the Kohonen layer, i.e., it accepts the target variables (t_{si}) of the target vector T_s in exactly the same way as the Kohonen layer accepts the input variables (x_{si}) of the input vector X_s, Figure 2.5, only that it does not influence to the positioning of objects.

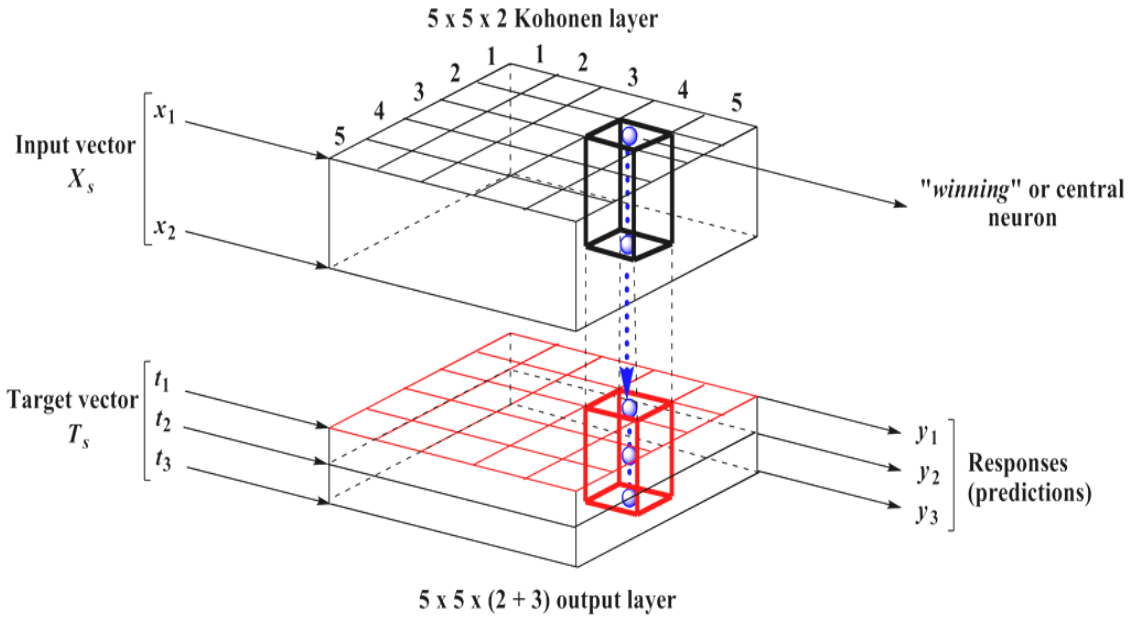


Figure 2.5. Counter-propagation artificial neural networks architecture (Saçan et al., 2013)

Compared to the KANN, where the learning procedure is performed in one-step, the learning algorithm in CP ANN is carried out in two steps. The first step (or unsupervised learning procedure) that corresponds to the mapping of objects onto the input layer (Kohonen layer) is exactly the same as the KANN learning procedure described previously, whereas the second step corresponds to a supervised learning procedure, i.e., for each input in the learning procedure, the target value is required. Consequently, the training of the network is possible with a set of input-target pairs {X_s, T_s}, where T_s designates the input vector (target vector) that carries the dependent variables (e.g., toxicity values). The output layer consists of different number of planes of weights as the target vector T_s has responses. In CP ANN, the correction of the weights (u_{ji}) in the output layer is performed according to the following equation (Eq. 2.7) (Heicht-Nielsen, 1987):

$$\Delta u_{j,i} = \eta\left(1-\frac{d}{p+1}\right)\left(x_{s,i}-u_{j,i}^{old}\right), d = 0,1,2,\ldots,p \qquad (2.7)$$

which is the same as (Eq. 2.5), except that in this case the weights (u_{ji}) are adapted to the target values (t_{si}), and not to the input variables (x_{si}) like in the Kohonen learning. The output layer is carrying the predictive capability of the CP ANN. Therefore, it can be successfully employed for the prediction of biological toxicity values not only for

existing/known compounds, but also for newly synthesized compounds (Saçan et al., 2013).

The CP ANN methodology has been successfully employed in past QSAR studies to discriminate between toxic modes of actions of phenols (Spycher et al., 2005) or to predict the toxicity of phenols to aquatic species (Novic and Vracko, 2003). Therefore, it is believed that CP ANN methodology should uncover successfully any non-linearity, if any, present in the algal toxicity data set generated in this study.

2.1.4. Validation of QSAR models

Due to the interest in the regulatory utilization of QSARs, there has been an increased awareness of the requirement for good practice in their application to make predictions of toxicity (Hewitt et al., 2007). The need for minimizing the subjectivity present in all stages of model development resulted in the so-called OECD principles for the validation of QSARs. To facilitate the consideration of a (Q)SAR model for regulatory purposes, the QSAR model should be associated with the following five categories of information: (i) a defined endpoint, (ii) an unambiguous algorithm, (iii) a defined domain of applicability, (iv) appropriate measures of goodness-of-fit, robustness, and predictivity, and (v) a mechanistic interpretation, if possible (OECD, 2007).

A defined endpoint signifies clarity in the endpoint being predicted by a given model. This is because a given endpoint can be determined by different experimental protocols and under different experimental conditions. It is, therefore, important to identify the experimental system that is being modeled by the QSAR (OECD, 2007).

An unambiguous algorithm is to ensure transparency in the model algorithm that generates predictions of an endpoint from information on chemical structure and/or physicochemical properties. This holds for the algorithm defining the QSAR model but also for the descriptors used in these algorithms (Zvinavashe et al., 2008).

Since a model is a limited description of reality, these limits should be well-defined (Worth and Cronin, 2004). The principle of applicability domain indicates that the QSARs

are associated with limitations regarding the type of chemical structures, physicochemical properties, and mechanism(s) for which a model generates reliable predictions. In the simplest of approaches, the domain of applicability can be defined by the boundaries of the descriptor values for the compounds in the training set (OECD, 2007; Zvinavashe et al., 2008). Any given QSAR is expected to make reliable predictions within its applicability domain (AD) for untested chemicals (Bhhatarai and Gramatica, 2010).

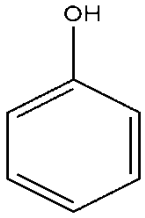
The fourth principle that requires appropriate measures of goodness-of-fit, robustness, and predictivity indicates that the parameters that reflect both the internal performance of a QSAR model and its predictivity should be provided. The internal performance is characterized by the goodness-of-fit and robustness of the model. The goodness-of-fit measures how well the model accounts for the variation in the response in the training set. The term robustness is used to measure the stability of the parameters and predictions when one or more of the training set chemicals are removed, and the model is regenerated excluding the removed compounds. The predictivity of a QSAR is usually characterized by the externally validated coefficient of determination (OECD, 2007; Zvinavashe et al., 2008). The real usefulness of a QSAR is not only its ability to reproduce known data verified by its fitting power (i.e., r^2), but its possibility of predictive application for untested chemicals (Gramatica et al., 2007). For this reason, recently much more emphasis has been put into the external validation of the QSAR models.

Finally, the OECD guidelines state that a QSAR model should be associated with a mechanistic interpretation (OECD, 2007). Such a mechanistic interpretation links the descriptors used in the model and the endpoint being predicted. This fifth principle also relates to the question as to how much can be learned from QSAR studies (Zvinavashe et al., 2008).

2.2. Phenols

In this thesis, phenols were selected as test chemicals because of their environmental and toxicological importance. The phenols selected for toxicological assessment in this thesis are listed in Table 2.4.

Hydroxybenzene, or phenol, is the parent molecule for the class of chemicals named phenols which carry the structure of a benzene ring with a hydroxyl group, as depicted in Figure 2.6.



Phenol

Figure 2.6. The parent phenol molecule

2.2.1. Environmental significance

Phenols have been in production since 1860's (Duchowicz et al., 2008). They have been used in the production of synthetic resins, dyes, pharmaceuticals, pesticides, synthetic tanning agents, perfumes, lubricating oils and solvents (Rayne et al., 2009). Owing to their widespread use, they have been detected in terrestrial and aquatic food chains (Jensen, 1996) and in environmental samples, particularly in those obtained from aquatic ecosystems (WHO, 1987, 1989, 1994).

Phenol is produced through both natural and anthropogenic processes. It is naturally occurring in some foods, in human and animal wastes, and in decomposing organic material. The largest use of phenol is as an intermediate in the production of phenolic resins, which are used in the plywood, adhesive, construction, automotive, and appliance industries. Phenol is also used in the production of synthetic fibers such as nylon and for epoxy resin precursors such as bisphenol-A. Because of its anesthetic effects, phenol is

used in medicines such as ointments, ear and nose drops, cold sore lotions, throat lozenges and sprays and antiseptic lotions (USEPA, 2002a).

Phenol and its simple halogenated derivatives are important industrial and natural products. Of the four halogenated classes of phenols (fluoro, chloro, bromo, and iodo), the chlorophenols have the highest industrial value (Rayne et al., 2009). The data set in this thesis includes the congeneric series of chlorophenols from mono-chlorophenols to pentachlorophenol (Table 2.4). These compounds represent a group of commercially produced substituted phenols used as intermediates in the synthesis of dyes, pigments, phenolic resins, pesticides and herbicides. Pentachlorophenol is widely used as a wood preservative (USEPA, 1980).

Chlorophenols are prepared by direct chlorination or the hydrolysis of higher chlorinated derivatives of benzene. Phenol has been reported to be highly reactive to chlorine which leads to the fact that chlorophenols may be inadvertently produced during the chlorination reactions for the disinfection of drinking water supplies. Pure chlorophenols are colorless crystalline solids, with the exception of 2-chlorophenol which is a clear liquid. As a group, the chlorophenols can be characterized by their unpleasant, pungent odor. They are weak acids; the acid dissociation constants (*p*Ka) for the chlorinated phenols decrease with increased substitution of chlorine atoms into the aromatic ring. The volatility of these compounds is also affected by the number of chlorine atoms present in the aromatic ring: the volatility of the chlorophenols decrease as the number of chlorine atoms increases (USEPA, 1980).

Due to the widespread use of chlorophenols and their persistence in the environment, human exposure has been demonstrated to occur regularly. For example, Mirabelli et al. (2000) found that both nasal and nasopharyngeal cancers were significantly associated with duration of occupational exposure to chlorophenols among men aged 30-60 in USA. The widespread exposure to chlorophenols has even led to the development of an immunoassay for 2,4,5-trichlorophenol which is one of the congeners with significant toxic effects (Nichkova et al., 2002).

In this thesis, apart from phenol and chlorophenols, 10 polyphenols (phenols carrying two or more hydroxyl groups) including catechols, resorcinols, hydroquinones and pyrogallol were also selected for toxicological assessment.

Table 2.4. The phenols selected for toxicological assessment

ID	Chemical	Abbreviation	CAS No
1	Phenol	P	108-95-2
2	2-chlorophenol	2-CP	95-57-8
3	3-chlorophenol	3-CP	108-43-0
4	4-chlorophenol	4-CP	106-48-9
5	2,3-dichlorophenol	2,3-DCP	576-24-9
6	2,4-dichlorophenol	2,4-DCP	120-83-2
7	2,5-dichlorophenol	2,5-DCP	583-78-8
8	2,6-dichlorophenol	2,6-DCP	87-65-0
9	3,4-dichlorophenol	3,4-DCP	95-77-2
10	3,5-dichlorophenol	3,5-DCP	591-35-5
11	2,3,4-trichlorophenol	2,3,4,-TCP	15950-66-0
12	2,3,5-trichlorophenol	2,3,5-TCP	933-78-8
13	2,3,6-trichlorophenol	2,3,6-TCP	933-75-5
14	2,4,5-trichlorophenol	2,4,5-TCP	95-95-4
15	2,4,6-trichlorophenol	2,4,6-TCP	88-06-2
16	3,4,5-trichlorophenol	3,4,5-TCP	609-19-8
17	2,3,4,5-tetrachlorophenol	2,3,4,5-TeCP	4901-51-3
18	2,3,4,6-tetrachlorophenol	2,3,4,6-TeCP	58-90-2
19	2,3,5,6-tetrachlorophenol	2,3,5,6-TeCP	935-95-5
20	Pentachlorophenol	PCP	87-86-5
21	1,2,3-trihydroxybenzene	Pyrogallol	87-66-1
22	Hydroquinone	Hq	123-31-9
23	Chlorohydroquinone	ClHq	615-67-8
24	Tetrachlorohydroquinone	TetraClHq	87-87-6
25	Catechol	Cat	120-80-9
26	4-chlorocatechol	4-ClCat	2138-22-9
27	3,5-dichlorocatechol	3,5-DiClCat	13673-92-2
28	Resorcinol	Res	108-46-3
29	4-chlororesorcinol	4-ClRes	95-88-5
30	4,6-dichlororesorcinol	4,6-DiClRes	137-19-9

Hydroquinone (1,4-dihydroxybenzene, Figure 2.7) is a white crystalline substance which is highly soluble in water. It is a reducing agent that is reversibly oxidized to semiquinone and quinone (WHO, 1996). Hydroquinone is naturally present in plants, most frequently as arbutin which is a glucoside (Varagnal, 1981, as cited in Topping et al., 2007). It also occurs in soils in combination with other phenols as polymeric phenolic substances, commonly referred to as humic acids, and in plant products such as coffee (Hogl, 1958, as cited in Topping et al., 2007). Besides its widespread occurrence in nature, hydroquinone is extensively used as a reducing agent, as a photographic developer, as an

antioxidant for many oxidizable products, as a stabilizer or polymerizing inhibitor for certain materials that polymerize in the presence of free radicals, and as a chemical intermediate for the production of antioxidants, antiozonants, agro-chemicals, and polymers. It is a skin lightening agent and is used in cosmetics, hair dyes, and medical preparations (WHO, 1996; Topping et al., 2007).

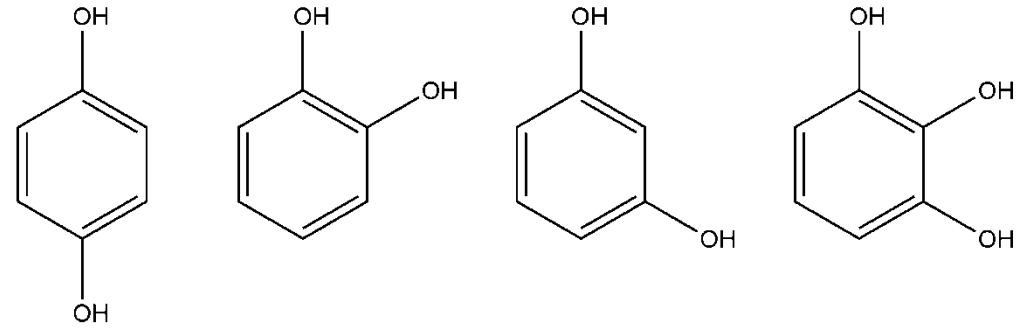


Figure 2.7. Chemical structure of hydroquinone, catechol, resorcinol and pyrogallol

Catechol (1,2-dihydroxybenzene, Figure 2.7) is used in pharmaceutical industry and as a reagent for photography, dyeing fur, rubber and plastic production (Schweigert et al., 2001a). Chloro-catechols, which are important chloro derivatives of the parent catechol molecule, are introduced into the environment by a variety of industrial and natural sources. These compounds are used as industrial reagents in the production of rubber, dyes, plastics, pharmaceuticals, and cosmetics, and they are important pollutants in the effluents of saw and pulp mills. Due to their resistance against conventional wastewater treatment processes, they are discharged into the receiving water bodies, and because of their hydrophobicity, they finally accumulate in sediments (Schweigert et al., 2001b).

Resorcinol (1,3-dihydroxybenzene, Figure 2.7) is a white crystalline compound with a weak odor and a bittersweet taste (WHO, 2006). It is released as a thermal breakdown product in wastewater effluents during coal-conversion processes (Upadhyay et al., 2010). The rubber industry is the largest user of resorcinol. Together with formaldehyde and synthetic rubber latex, it is used in the manufacture of tyres for passenger cars, trucks, off-road equipment, and other fibre-reinforced rubber mechanical goods, such as conveyor and

driving belts. It is also used for high-quality wood bonding applications. Other uses include the manufacture of dyestuffs, pharmaceuticals, flame retardants, agricultural chemicals, fungicidal creams and lotions, explosive primers, antioxidants, a chain extender for urethane elastomers, and a treatment to improve the mechanical and chemical resistance of paper machine fabrics (WHO, 2006).

Among the polyphenols used in this thesis, pyrogallol (1,2,3-trihydroxybenzene, Figure 2.7) is the only compound that has three hydroxyl groups attached to the benzene ring. Pyrogallol is widely distributed in nature, as it is used widely in many industries and consumer products. It is naturally present in oak, eucalyptus and other hardwood plants, as a decomposition product of hydrolysable tannins and possesses anti-fungal and anti-psoriatic properties. In natural conditions, pyrogallol is present as a contaminant in tannins, anthocyanins, flavones and alkaloids and released into environment during its isolation, disposal and industrial use. Due to its oxygen radical generating property, pyrogallol is commonly used as a photographic developing agent, in hair dying industry and as an antiseptic (Upadhyay et al., 2010).

2.2.2. Mode of action of phenols

Mode of action[2] (MOA) is defined as a drug or chemical action referring to the type of response produced in an exposed organism (Borgert et al., 2004). The biological membranes are among the most important target sites for toxic effects. Partitioning of a hydrophobic compound into membranes causes disturbances in the structure and functioning of the membranes leading to the so-called baseline toxicity, or narcosis, which constitutes the minimal toxicity of any hydrophobic pollutant (Escher and Schwarzenbach, 2002). Narcosis (i.e., non-polar and polar narcosis) is the least specific, but arguably the most important mode of toxic action in ecotoxicology because approximately 70% of all organic industrial chemicals are believed to act via narcosis (Bradbury and Lipnick, 1990). There are several hypotheses on how the molecular mechanism of narcosis functions which has been excellently reviewed by Escher and Schwarzenbach (2002). In summary, baseline

[2] The term mode of action should not be confused with another closely related term, namely the mechanism of action, which implies a more detailed understanding and description of events, often at the molecular level, than is meant by mode of action (USEPA, 2005).

toxicity has been regarded as a disruption of membrane functions, although the detailed mechanism remains unclear (Öberg, 2004).

Several approaches have been developed to classify chemicals based on their MOAs usually placing narcosis as a base for classification. Verhaar et al. (1992) introduced a classification scheme for MOA based on the so-called toxic ratio (TR) concept which was defined as the ratio of the effect concentration (EC_x) of a given endpoint predicted for baseline toxicity ($EC_{x,}$ baseline) to the experimentally determined effect concentration ($EC_{x,}$ experimental) (Eq. 2.8):

$$TR = EC_x, \text{ baseline} / EC_x, \text{ experimental} \quad (2.8)$$

According to this scheme, a compound was classified as a polar narcotic if it was five to ten times more toxic than would be predicted from a QSAR of non-polar narcosis ($5 \leq TR \leq 10$). Compounds with a $TR \geq 10$ were classified as reactive or specifically acting (Verhaar et al., 1992). The group of specifically acting compounds was not further detailed in that study. If a specifically acting compound had a $pKa \leq 6.5$, it was classified as an uncoupler (Schultz, 1987). Uncouplers have been defined as a compound that interferes reversibly with mitochondrial membranes and with the build-up of the proton gradient (Aptula et al., 2005b). To elucidate the mechanisms of uncoupling, Escher and coworkers (1999) have established a novel test system, called Kinspec. Briefly, the target site of the effect, the so-called “energy-transducing membrane”, is isolated in this test system and the measured effects can be directly related to concentrations in the membrane with appropriate membrane-water partitioning experiments (Escher et al., 2000a; Escher et al., 2000b). At the end of this experimental procedure, the measurement of the intrinsic toxicity (toxicodynamic phase) can be separated clearly from the toxicokinetic phase, i.e., uptake, partitioning within the organism, metabolism, and excretion (Escher and Schwarzenbach, 2002).

It should be emphasized that the MOA concept is a controversial topic and often a definitive classification of chemicals is not possible due to the differences in scientific approaches. For example, it was proposed that chemically inert phenols such as chlorophenols exhibit toxicity by membrane perturbations, a mechanism commonly

referred to as polar narcosis (Aptula et al., 2005b). Based on this MOA assignment, several QSAR studies have treated these compounds (except tetra-chloro substituted phenols and pentachlorophenol which were classified as uncouplers) as polar narcotics (Cronin et al., 2002; Enoch et al., 2008). On the other hand, as briefly introduced above, Escher and Schwarzenbach (2002) classified majority of chlorophenols as respiratory uncouplers with varying potencies, based on their findings in the Kinspec test system. From a purely statistical perspective, Schüürman et al. (1997) indicated that the uncoupling activity of pentachlorophenol, which has been regarded as a classical uncoupler, did not significantly increase its overall acute toxicity over polar narcosis in ten different endpoints. As can be seen from this mini-review of phenolic MOA assignments based on different methods, it is often not possible to reach a consensus opinion on the MOA even for relatively well-studied chemicals, such as chlorophenols. On the other hand, it is possible to state that phenols are associated with a wide range of possible MOAs, ranging from narcotic to electrophilic (reactive) mechanisms (Aptula et al., 2005b). Among these, the reactive mechanisms are more difficult to delineate (Aptula et al., 2005b), because the so-called reactivity is not specific to a certain reaction; instead, it includes a number of competing processes and different chemical reactivity mechanisms (Schultz, 1999). The majority of effects attributed to reactive toxicity arise from the reaction between an electrophilic chemical and a biological nucleophile (Schultz et al., 2006).

For regulatory usage, it is important to construct and evaluate models based on MOA (i.e., local models) because this enables confidence to be better assigned to the predictions when these models are used in a real life scenario (Enoch et al., 2008). However, as discussed above, it should be kept in mind that the knowledge on the MOA of chemicals varies considerably (Spycher et al., 2008) and assignment of the correct MOA is a difficult task (Schultz, 1999). As stated by Aptula et al. (2005b), certain mechanisms may even be specific for the given species and these can be related to the metabolism and distribution of the compounds *in vivo*. Therefore, the models that accurately predict the toxicity of chemicals without first identifying their toxic mechanisms (i.e., global models) are highly desirable (Schultz, 1999). In this thesis, the QSARs were initially developed without taking MOA of phenols into consideration. However, once a robust, valid and predictive QSAR model is constructed, it was investigated on the basis of the mechanistic make-up of the data set. This led to the development of several local models (i.e., models based on MOA).

This approach proved to be very useful in uncovering the latency present in the data set as well as in establishing a mechanistic interpretation of the models to be evaluated from a risk assessment perspective.

2.3. Risk assessment perspective

Environmental risk assessment stands in the intersection where environmental science and risk management meet. It aims to provide adequate information to decision makers with the aim of protecting the environment from adverse effects of chemicals (Breitholtz et al., 2006). However, the risk assessment process has been severely criticized for being both too slow and too simplistic in terms of scientific approach. The pressure from various stakeholders to replace the scientific approach resulted in the so-called "precautionary principle" which suggests applying controls/restrictions to chemicals in advance of scientific understanding if there is a presumption that harm will be caused (Forbes and Calow, 2002). In the EU, risk assessment proceeds along the lines of the precautionary principle which mandates precautionary action until the uncertainties can be reduced if they are too large (Hommen et al., 2010). The fact that the so-called precautionary principle has superseded the "scientific approach" highlights the dilemma of modern times, the dilemma of balancing the benefits of industrial output (e.g., products such as cosmetics, textiles, etc.) with environmental protection. Considering the ever increasing number of chemicals that must be evaluated for their potential hazards and risks for the environment, the evaluation process seriously challenges the ability of the scientific approach to properly inform the assessors (Forbes and Calow, 2002).

As mentioned previously, the REACH legislation aims to control the risks associated with the manufacture/import/usage for all chemicals in the EU at a tonnage level above one ton per year. To gather relevant information on the hazard of substances, data obtained from algal growth inhibition tests are mandatory in REACH. However, considering the ever increasing number of chemicals, the implementation of REACH requires new methods to allow the risk assessment of a vast amount of substances before May 31st, 2018 (Vink et al., 2010). To this end, REACH provides a flexible framework for the use of non-testing methods including QSAR models provided that a) the model is scientifically valid, b) the model is applicable to the chemical of interest, c) the prediction is relevant for the

regulatory purpose and d) appropriate documentation on the method and result is provided (Worth, 2010).

From a risk assessment perspective, there are three issues that are particularly important in this thesis. The first one is the requirement of algal growth inhibition tests in REACH legislation. The other is the use of QSAR methodology to provide information on chemicals that has not been tested for their toxic effects. Another issue that has been largely overlooked is the lack of data for marine species. This might probably stem from the relatively new recognition of the concerns regarding the environmental status of marine environments. Indeed, this official recognition dates back to only 20 years ago when the protection of the marine environment was recognized at the Earth Summit in Rio de Janeiro in 1992 following the adoption of Agenda 21, a programme for sustainable worldwide development. In that meeting an exclusive section was devoted to protection of the oceans and seas, and particular concern regarding pollutants which exhibit toxicity, persistence and bioaccumulation was expressed for the first time in an international scale (UN, 1992). It called on governments to take priority action to control such substances and promote risk and environmental impact assessment to ensure an acceptable level of environmental quality (ECETOC, 2001). Yet, the great majority of the directives that have been developed in the European Union focus mainly on freshwater environments. Other environmental compartments, including marine ecosystems, are discussed as well, but they are explicitly considered in these directives (Hommen et al., 2010). The development of directives focusing on the protection of freshwater ecosystems has been a result of urbanization which has rendered freshwater ecosystems more vulnerable to the chemical releases and this made the protection of freshwater communities the first priority by regulatory schemes. Consequently, there is now substantial amount of information available on the toxicity of chemicals to freshwater organisms while there are relatively fewer data on the effect of chemicals to marine organisms (ECETOC, 2001), in particular to aquatic plants and algae (Hutchinson et al., 1998). Therefore, this thesis was structured so as to investigate the possible use of alternative methods, particularly the QSAR methodology, in efforts to fill the data gap present in algal toxicity data, especially for marine algae. Apart from the QSTR models developed using marine algal toxicity data, other approaches, such as the use of data from other test systems *in lieu* of marine algal data was also explored in detail.

2.3.1. Interspecies toxicity correlations

The traditional use of QSAR in environmental risk assessment is to predict the toxicity of chemicals not included in the training set. Instead of prediction, there are also studies where QSAR methodology is used for the investigation of reactive mechanisms (Aptula et al., 2005b) or discrimination of MOAs (Schüürman et al., 1997; Yao et al., 2005). An interesting use of QSAR methodology, which has received little attention, is its potential to be used as a tool to predict the toxicity of a compound towards a particular biological organism using toxicity data on another species (Zvinavashe et al., 2008). Interspecies toxicity correlations (or interspecies toxicity relationships) have the potential to be used as a tool for estimating contaminant sensitivity with known levels of uncertainty for a diversity of different species (Raimondo et al., 2007; Raimondo et al., 2010). In fact, the possibilities for these extrapolations have been demonstrated previously in the literature for industrial chemicals (Kaiser, 1998; Dimitrov et al., 2000; Zhang et al., 2010), even for pharmaceuticals (Kar and Roy, 2010). However, in these studies, not much attention has been directed to algae.

Within the scope of this thesis, several interspecies toxicity correlations between algae and other aquatic organisms were developed in order to a) predict the toxicity of untested chemicals towards algae, b) gain insights to the toxicity of phenols towards different aquatic organisms and c) use freshwater data *in lieu* of marine algal toxicity data. For the latter, it should be indicated that the use of freshwater data in marine hazard assessments is described in the European Union Technical Guidance Document (EU TGD, 2003), Canadian Water Quality Guidelines (CCME, 2007) and the USA marine water quality criteria (Russo, 2002). In using this surrogate approach, the EU TGD (2003) calls for a clear understanding of the comparability of effects data generated on both types of species and require argumentation on a case-by-case basis for the inclusion of freshwater data in marine hazard assessment. As suggested in the ECETOC workshop (2008) on the probabilistic approaches for marine risk assessment, the MOA approach can be useful to analyze aquatic toxicity data and extract useful information for decision making. Therefore, the interspecies toxicity correlations developed in this thesis were also evaluated based on the MOA approach to be able to generate information compatible with the requirements of regulatory schemes.

As mentioned above, the interspecies toxicity correlations have the potential to fill the data gaps in marine algal ecotoxicity data, if statistically significant and mechanistically interpretable relationships can be demonstrated. To this end, several aquatic species (bacterium, protozoan, alga, daphnid and fish) were selected for interspecies toxicity correlations based on the ecological significance of the organisms as well as the availability of toxicity data on the selected species.

2.3.1.1. Bacteria (*Vibrio fischeri*). *Vibrio fischeri* (formerly known as *Photobacterium phosphoreum*) is a gram-negative bacterium found globally in marine environments (Madigan and Martinko, 2005). The photo luminescent bioassay uses a suspension of *V. fischeri* bacteria in saline water (to protect bacteria from osmotic damage) and measures the reduction in light output of its natural luminescence on exposure to the toxicant (Kaiser, 1998). Since this bacterial bioassay can be performed with relative ease at a limited cost, extensive toxicity data for *V. fischeri* has been generated and are available in the form of monographs (Kaiser and Palabrica, 1991) or in commercial databases (TerraTox™, 2006) covering over 1700 chemicals.

Due to the availability of a huge toxicity database on *V. fischeri*, it has been of particular interest to conduct interspecies toxicity correlation between the marine bacterium and other aquatic species. In this respect, one of the earliest efforts to exploit huge amount *V. fischeri* toxicity data was attempted by Kaiser (1998), who undertook an extensive study to establish correlations with data from other test systems including algae, daphnia, fish and other aquatic and terrestrial organisms. Of particular importance was the relationship between the toxicity of 59 compounds to *V. fischeri* and freshwater alga *Chlorella pyrenoidosa* which resulted in a promising correlation (n=59, r^2=0.76). On the other hand, in a more recent study, Zhang et al. (2010) reported a rather poor relationship between *V. fischeri* and freshwater alga *Scenedesmus obliquus* (n=30, r^2=0.66). The disparity of the two correlations between bacteria and algae in the examples above suggests that the significance of the interspecies relationship might vary for each algal test system. The toxicity correlation between two marine species might give better results compared to those between a marine and a freshwater organism (Sverdrup et al., 2002); as such, the interspecies toxicity correlation with marine alga *D. tertiolecta* and *V. fischeri* merits special investigation because both species are marine dwelling organisms. In

addition, the interspecies toxicity correlation between *V. fischeri* and freshwater alga *C. vulgaris* provides the opportunity to increase the knowledgebase on the relationships between bacteria and freshwater algae.

2.3.1.2. Protozoa (*Tetrahymena pyriformis*). *Tetrahymena pyriformis* is a ubiquitous ciliated protozoan. It can be cultured and maintained in a small volume of culture medium. Its generation time is fast, so the effects of tested variables can be studied through several generations (Lukacinova et al., 2007). *T. pyriformis* has been tested on a large number of organic chemicals. The TETRATOX database (2012) includes comprehensive toxicity results for approximately 2400 aliphatic and aromatic chemicals. This database represents an invaluable source for toxicity data that has been used in QSAR modeling (Devillers, 2008). Considering the availability of large database of acute toxicity data on *T. pyriformis*, an interspecies toxicity correlation study between algae and protozoa is of considerable interest to predict the toxicity of chemicals available in the protozoan data set but not to algae.

Previous studies demonstrated the possibility of constructing reliable interspecies toxicity correlations between *T. pyriformis* and other organisms. For example, Cronin and Schultz (1998) found an excellent correlation between *T. pyriformis* and *V. fischeri* for ketones, aldehydes and α,β-unsaturates. In another study, Seward et al. (2002) tested the correlation between the toxicity of a heterogeneous set of chemicals to *T. pyriformis* and *P. promelas*, concluding that for phenols, there was a strong correlation between these two endpoints (n=29, r^2=0.81). However, an interspecies toxicity correlation between algae and *T. pyriformis* is limited in the literature. In a recent study where they investigated the interspecies correlations of toxicity to eight aquatic species, Zhang et al. (2010) did not find a significant relationship between algae and protozoa. In another study, Cronin et al. (2004) demonstrated a strong correlation between *T. pyriformis* and *C. vulgaris* using a 15-minute test methodology to determine the algal toxicity of chemicals. As discussed previously, the 15-min bioassay is not based on an endpoint compatible with standardized algal test protocols; therefore, the interspecies toxicity correlation between *T. pyriformis* and *C. vulgaris* should be verified with toxicity data generated according to established protocols, such as the one generated in this thesis.

2.3.1.3. Algae. As demonstrated by Aruoja et al. (2011), algal QSARs are remarkably low, possibly due to the lack of reliable algal data in the literature (Cronin et al., 2004). The novel toxicity data generated in this thesis using marine and freshwater algae provides the unique opportunity to investigate inter-algal toxicity correlations using data generated in the same laboratory. Additionally, the data set reported by Aruoja et al. (2011) comprising the 72 h growth inhibition data (pEC_{50}, in mmol/L) of 58 compounds towards freshwater alga *P. subcapitata* was used as an external validation set to assess the predictive power of the algal QSAR models developed in this thesis. It should be noted that *P. subcapitata* is a freshwater algal species conducted routinely in ecotoxicity testing (Mayer et al., 1998; Christensen et al., 2009; Aruoja et al., 2011). The external data set was generated in the same laboratory, by the same researcher following standardized algal growth inhibition test protocols; therefore, it should provide a suitable basis upon which to evaluate the external predictivity of the algal QSARs constructed in this study. Accordingly, the algal QSARs as well as the interspecies toxicity correlations between *D. tertiolecta*, *C. vulgaris* and *P. subcapitata* should provide a preliminary analysis as to the possibility between extrapolating between different algal species.

In their critical review of the state of the art of marine risk assessment, Peters et al. (2005) conducted an extensive study to compare the responses of marine and freshwater organisms to chemicals. The researchers noted that for crustaceans, algae and further taxonomic groups the number of marine toxicity data were limited. The available data on the freshwater and saltwater toxicity of algae (n=8) showed no functional relationships. However, as also stated by the researchers, their study on the relative sensitivity of freshwater and marine organisms suffered from a high level of background-variability because the data was not produced for such a comparison. In contrast, the data generated in this study provides the unique opportunity to compare the responses of marine and freshwater algae, because the data was generated for such a comparison; therefore should provide a realistic assessment with minimum background variability.

2.3.1.4. Daphnia (*Daphnia magna*). *Daphnia magna* is a small cladoceran crustacean which is one of the most widely used animals in ecotoxicity testing (Billoir et al., 2007). They play a crucial role in the sustainability of ecosystems because they are the primary consumers in aquatic food chains. Since algae are the primary producers in aquatic

ecosystems, an interspecies toxicity correlation between algae and daphnia is particularly interesting between these two species considering the interactions between daphnia and algae in aquatic ecosystems.

Previous attempts to construct interspecies toxicity correlations using data on *D. magna* and algae resulted in rather poor relationships. For example, Zhang et al. (2010) reported an unsatisfactory correlation between *S. obliquus* and *D. magna* (n=37, r^2=0.32). In another study, Sverdrup et al. (2002) found that for 30 exploration and production chemicals, the interspecies toxicity correlation between *D. magna* and marine diatom alga *Skeletonema costatum* was rather low (n=30, r^2=0.61). In their analysis of toxicity data sets provided by the Japanese Ministry of Environment, Tebby et al. (2011) found that the correlation between the toxicity of chemicals to alga *P. subcapitata* and *D. magna* was low (n=85, r^2=0.54). In another study, based on their analysis on the data sets provided by the French Ministry of Ecology and Sustainable Development, Henegar et al. (2011) also reported a rather low correlation (n=164, r^2=0.41) between algae (toxicity data on *P. subcapitata* and *Desmodesmus subspicatus* previously known as *Scenedesmus subspicatus*) and *Daphnia magna*. Zhang et al. (2010) suggested that the reason behind the low correlations observed between *D. magna* and algal toxicity data are due to the differences in bio-uptake and mode of action of chemicals. No matter what the underlying reason is, the brief literature survey points towards rather unreliable correlations between algae and daphnia. Within the scope of this thesis, the toxicity data on two algae (*D. tertiolecta* and *C. vulgaris*) and *D. magna* will be correlated in an effort to improve our understanding of the toxicity correlations between algae and daphnia.

2.3.1.5. Fish (*Pimephales promelas*). Fish occupy the highest trophic level in most aquatic ecosystems. The REACH legislation requires short-term toxicity data on fish for compounds manufactured or imported > 10 tpa (EC, 2006). Together with algae and daphnia, eco-toxicological information is required for fish to be able to develop a sound risk assessment framework for the protection of fragile aquatic food chains. Therefore, similar to the relationship between algae and daphnia, a sound interspecies toxicity relationship between algae and fish can also prove very useful to evaluate the risks of a chemical on aquatic food chains. Moreover, such a relationship might lead to the exploitation of high quality fish toxicity databases such as the one that has been developed

by the United States Environmental Protection Agency Duluth Laboratory that contains 96 h LC_{50} data on *Pimephales promelas* (fathead minnow) (Russom et al., 1997).

In their analysis of the acute toxicity data gathered by the Japanese Ministry of Environment, Tebby et al. (2011) recently reported that the correlation between the toxicity of chemicals to alga *P. subcapitata* and the fish species *Oryzias latipes* was very low (n=85, r^2=0.41). The authors also reported that the interspecies correlations between fish and daphnia toxicity were higher than their correlation with algae. In another study, Henegar et al. (2011) analyzed the database provided by the French Ministry of Ecology and Sustainable Development and also found that daphnia and fish experimental data had better correlation among each other compared to the correlations between daphnia-algae and fish-algae. The authors reported that fish toxicity was able to explain 52% of the variance in algal toxicity of 185 chemicals. It should be noted that the statistical analysis carried out by Henegar et al. (2011) were based on a data collection using toxicity values determined for two different algal species (*P. subcapitata* and *D. subspicatus*) and three different fish species (*Danio rerio*, *Cyprinus carpio* and *Oncorhynchus mykiss*). The use of different species in the toxicity correlations or treating different species as a single organism (i.e., treating *P. subcapitata* and *D. subspicatus* as a single organism) can be expected to introduce additional variance in the analysis. For example, in the ring test participated by 18 laboratories, ISO (2004) reported that the EC_{50} of 3,5-dichlorophenol, a well-known reference toxicant, was found to be 6.42±0.56 mg/L (mean±std.error) for *D. subspicatus* while the EC_{50} value of the same chemical to *P. subcapitata* was 3.38±0.43 mg/L. As can be seen from this example, the toxicity of a well-known polar narcotic chemical varies approximately two orders of magnitude between different algal species. Therefore, an interspecies toxicity correlation should ideally be conducted using data on single species. As such, this thesis provides the unique opportunity to test the relationship between fish and algae using data on single species which would not pose any additional variance in the analysis other than the biological data.

2.3.2. ECOlogical Structure-Activity Relationship Model (ECOSAR)

Due to the large number of chemicals that needs to be evaluated from a risk assessment perspective, industries, regulatory agencies and non-governmental

organizations have a growing interest in reliable QSAR models (ECETOC, 2003; Sanderson et al., 2003; Worth, 2010). The most extensively validated and used QSAR is the ECOSAR (ECOlogical Structure-Activity Relationship Model**)** which is an easy-to-use computer programme developed and routinely applied by the United States Environmental Protection Agency (USEPA) for predicting aquatic toxicity to fish, aquatic invertebrates (daphnids), and green algae (Reuschenbach et al., 2008). The latest version of the programme is freely available from the USEPA (2011).

SARs for more than 50 different chemical classes are integrated in the ECOSAR programme. The SARs are all based on algorithms reflecting a linear regression relationship between the decadic logarithm of the aquatic toxicity of the chemicals used to develop the SAR for the specific chemical class (training set) and Log K_{ow}. ECOSAR contains a library of class-based QSARs for predicting aquatic toxicity, overlaid with an expert decision tree for selecting the appropriate chemical class. ECOSAR version 1.1 is programmed to identify 111 chemical classes and allows access to 709 QSARs for numerous endpoints and organisms (ECOSAR, 2011).

Although the ECOSAR predictions do not replace the need for experimental determination of algal toxicity, they can be used to estimate potential hazards of chemicals in the environment where there is no available data. Therefore, the toxicity data generated in this study was also compared to the predictions obtained from ECOSAR for green algae.

3. MATERIALS AND METHODS

3.1. Test chemicals

All phenols used in this thesis for toxicological assessment were purchased with the financial support of Boğaziçi University Research Funds (Project No: 09Y105P and 5564) from Sigma-Aldrich Co., except phenol which was obtained from Merck Co. The majority of the chemicals had purity ≥98%; therefore, no further purification was undertaken.

For the tests carried out using marine alga, the stock solutions were prepared below the water solubility limits of each compound using artificial seawater which was prepared according to standard methods (APHA-AWWA-WEF, 1998). As for the tests carried out using freshwater alga, the stock solutions were prepared below the water solubility limits of each compound using de-ionized water. The stock solutions of tetrachlorophenols, pentachlorophenol, chlorohydroquinone and tetrachlorohydroquinone were prepared in dimethyl sulfoxide (DMSO). For the tests using these compounds, an additional solvent control containing the maximum DMSO concentration (0.1% v/v) was employed. Statistical analysis using *t*-test revealed no significant difference (p>0.05) between the growth of both algae in controls with and without DMSO. The inhibitory concentration of these chemicals was calculated taking the growth in solvent controls into account.

3.2. Experimental methods

3.2.1. Analytical methods

The analytical methods used in the experiments followed the Standard Methods for the Examination of Water and Wastewater (APHA-AWWA-WEF, 1998) except heavy metal, sulfate and gas chromatographic measurements. The methods used to quantify these measurements are explained below:

a) *p*H, salinity and conductivity analysis: The *p*H of seawater and the growth medium containing the control cultures of each bioassay were measured with a *p*H-meter (WTW

*p*H 330i) using a special electrode for saline water (WTW *p*H-electrode Sen Tix HW). The salinity and conductivity of seawater samples were measured with a conductivity meter (WTW LF320/Set).

b) Phosphate analysis: Ascorbic acid method (method no: 4500B, 4-146, APHA-AWWA-WEF, 1998) was used to quantify phosphate concentrations in seawater.

c) Nitrate analysis: Ultraviolet spectrophotometric screening method (method no: 4500B, 4-115, APHA-AWWA-WEF, 1998) was used to quantify nitrate concentrations in seawater.

d) Sulfate analysis: Sulfate concentrations in seawater were measured using ion chromatograph (Dionex, ICS-3000).

e) Chloride analysis: Chloride concentrations were determined according to the argentometric method (method no: 4500B, 4-67, APHA-AWWA-WEF, 1998).

f) Alkalinity analysis: Alkalinity of seawater was determined according to the titrimetric method (method no: 2320B, 2-27, APHA-AWWA-WEF, 1998).

g) Heavy metal analysis: Concentrations of aluminum (Al), lead (Pb), zinc (Zn), cadmium (Cd), copper (Cu), chromium (Cr), nickel (Ni) and cobalt (Co) were analyzed using Inductive Coupled Plasma (Perkin Elmer, Optima 2100 DV).

h) Gas chromatographic analysis: Actual concentrations of each chemical were measured by gas chromatography (Agilent, 6890N equipped with an automatic sampler, split/splitless injection port and flame ionization detector) at the beginning of the experiments. The actual concentrations were verified using a calibration curve of at least four concentrations prepared in methylene chloride. The maximum test concentrations or stock solutions of phenols were also analyzed in a separate test vessel without algae at the beginning and at the end of the experiments to check if there was a significant chemical loss due to volatilization, adsorption on the test vessel, etc. during the experiment. One mL sample was extracted in 0.5 mL methylene chloride in 2 mL vials capped with Teflon-lined

septum caps. The injector and detector temperatures (250^{o}C and 300^{o}C, respectively) were held constant during the analysis. HP-5MS capillary column used for separation was 30-m long, and had 0.25-mm inner diameter and 0.25-μm film thickness. The oven was programmed for an initial temperature of 40^{o}C for 1 min and then increased to 140^{o}C for 10^{o}C/min and then to 260^{o}C for 20^{o}C/min. Helium was used as the carrier gas at a constant flow rate of 33.3 cm/s and the injector was operated in splitless mode.

3.2.2. Algal growth inhibition assays

Algal growth inhibition tests were performed in batch cultures according to the standard procedures (APHA-AWWA-WEF, 1998; OECD, 2006) using the marine alga *Dunaliella tertiolecta* and freshwater alga *Chlorella vulgaris*. Parent cultures of *Dunaliella tertiolecta* had been obtained from TÜBİTAK/MAM and they have been maintained in the laboratory conditions for many years by Prof. Melek Türker Saçan for her past studies. The parent cultures of freshwater algae, *Chlorella vulgaris* strain (CCAP 211/11B) was purchased from Culture Collection of Algae and Protozoa – (CCAP, The Scottish Association for Marine Science, Scottish Marine Institute, Dunbeg, Argyll, UK) for this thesis with the financial support of Bogazici University Research Funds (Project No: 5564). All tests were conducted in a laminar air flow cabinet reserved for microbiological assays, which was pre-sterilized with ultraviolet light for at least an hour (Figure 3.1).

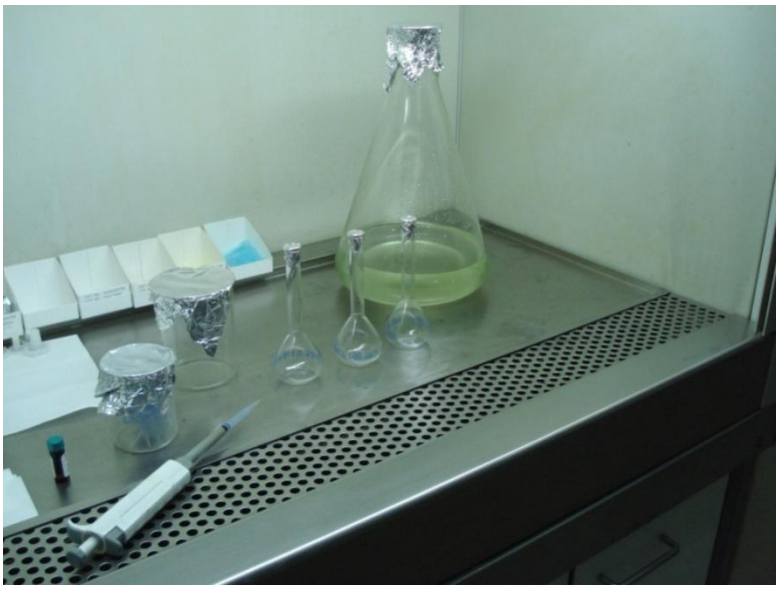


Figure 3.1. Photographic view of the laminar air flow cabinet

The growth medium used in the experiments for marine and freshwater algae is provided in Table 3.1 and Table 3.2, respectively.

Table 3.1. Bold basal medium with 3-fold nitrogen and vitamins used for bioassays conducted with freshwater algae, *Chlorella vulgaris*

Chemical	**Concentration (mg/L)**
$NaNO_3$	750
$CaCl_2.2H_2O$	25
$MgSO_4.7H_2O$	75
$K_2HPO_4.3H_2O$	75
KH_2PO_4	175
NaCl	25
$FeCl_3.6H_2O$	0.97
$MnCl_2.4H_2O$	0.41
$ZnCl_2$	0.05
$CoCl_2.6H_2O$	0.02
$Na_2MoO_4.2H_2O$	0.04
Na_2EDTA	7.5
Thiaminehydrochloride	12
Cyanocobalamin	10

Table 3.2. Modified f/2 growth medium used for bioassays conducted with marine algae, *Dunaliella tertiolecta*

Chemical	**Concentration (mg/L)**
$NaNO_3$	75
$NaH_2PO_4.2H_2O$	5
$Na_2SiO_3.H_2O$	12.9
$CuSO_4.5H_2O$	0.01
$ZnSO_4.7H_2O$	0.022
$MnCl_2.2H_2O$	0.18
$CoCl_2.6H_2O$	0.01
$FeCl_3.6H_2O$	3.15
$Na_2Mo_4.2H_2O$	0.06

The marine algal bioassays were carried out in filtered (GF/C Glass microfiber Whatman filters, England) seawater and enriched with modified f/2 medium (Okay and Gaines, 1996). Seawater was taken from the Sea of Marmara, near the coast of Samatya in Istanbul and stored in a freezer at –24^{o}C in a plastic container after filtration. Natural seawater characterization (Table 3.3) was made based on the standard procedures (APHA-AWWA-WEF, 1998). Additionally, the concentration of environmentally significant heavy

metals (Table 3.3) in the seawater was measured using Inductive Coupled Plasma (Perkin Elmer, Optima 2100 DV).

Table 3.3. Natural seawater characterization

Parameter	**Measurement**	**Metal**	**Concentration (µg/L)**
*p*H	8.3 ± 0.1	Al	46 ± 12
Salinity	23 ± 3 ‰	Pb	BDL[a]
Conductivity	30 ± 2 mS/cm	Zn	27 ± 26
Chloride	11.5 ± 0.9 g/L	Cd	BDL
Alkalinity	150.3 ± 5.1 mg/L $CaCO_3$	Cu	6 ± 2
Nitrate	0.44 ± 0.1 mg/L	Cr	2 ± 1
Phosphate	0.71 ± 0.2 mg/L	Ni	4 ± 3
Sulfate	1200 ± 60 mg/L	Co	BDL

[a] BDL: Below Detection Limit

Experiments were carried out using pre-sterilized equipment. The sterilization of glassware was carried out in a temperature controlled oven (WiseVen Fuzzy Control System, South Korea) at 180˚C for 2 hours. The plastic equipment (pipette, magnetic stirrers, etc.) and algal growth medium were autoclaved at 120^{o}C and under 2 atm. for 15 minutes.

All the glassware used during the experiments were cleaned with dilute nitric acid (%10 v/v) and then washed three times with tap water. Then, hexane was used to remove possible organic content in the glassware (remnants of toxicants). Again the glassware was washed with tap water three times rigorously and finally rinsed three more times with distilled water. After each use, the spectrophotometer cuvettes were also cleaned with hexane, washed three times with distilled water and left for drying for 1h.

For the marine algal toxicity assays, the inoculum was prepared with algae harvested from four to five day old cultures in exponential growth phase. Each milliliter of inoculum contained approximately 10^4 cells. Experiments were carried out in a temperature controlled growth chamber (18 ± 0.5^{o}C). Continuous illumination (30 µmol photons m^{-2} s^{-1} - at the level of test solutions) (APHA-AWWA-WEF, 1998) was provided from a rack of cool-white fluorescent tubes, horizontally arranged above a light reflecting platform where the test vessels were located (Figure 3.2).

Figure 3.2. Photographic view of the growth chamber

For the toxicity assays employing freshwater alga *Chlorella vulgaris* as the test organism, the inoculum was prepared with algae harvested from five d old cultures in exponential growth phase. Each milliliter of inoculum contained approximately $1.5x10^5$ cells. Experiments were carried out in the temperature controlled growth chamber (24 ± 0.5^oC) under continuous illumination (60 µmol photons m^{-2} s^{-1} - at the level of test solutions).

Following range finding assays, definitive experiments were carried out in three replicates using five concentrations of the test chemical. 100 mL test medium with algae was dispensed into sterile 500 mL borosilicate Erlenmayer flasks. For the marine algal tests, autoclaved magnetic stirrers were added to the test vessels to prevent algal attachment to the glass surface, and each vessel was stirred gently prior to sampling. The test vessels were shaken daily by hand during all experiments. The test conditions and other relevant information for the two algal assays are summarized in Table 3.4.

Table 3.4. Test conditions for the marine and freshwater bioassays

	Marine bioassay	Freshwater bioassay
Test type	Static non-renewal	Static non-renewal
Test organism	*Dunaliella tertioleca*	*Chlorella vulgaris*
Starting inoculum	10^4/mL	$1.5x10^5$/mL
Temperature	18 ± 0.5 ^oC	24 ± 0.5 ^oC
Light quality	Cool white fluorescent lighting	Cool white fluorescent lighting
Light intensity	30 μmol photons m^{-2} s^{-1}	60 μmol photons m^{-2} s^{-1}
Photoperiod	Continuous illumination	Continuous illumination
Test chamber size	500 mL	500 mL
Test solution volume	100 mL	100 mL
Replicates	3	3
Agitation	Daily by hand	Daily by hand
Test concentrations	Five and a control	Five and a control
Test duration	96 h	96 h
Endpoint	Growth	Growth

The algal growth inhibition test guidelines prepared by the Organization for Economic Cooperation and Development (OECD, 2006) recommends that at the end of 72 h, the algal population in the controls should increase by at least a factor of 16. However, as stated in the guidelines, this criterion may not be met when species that grow slower are used. In this case, the test period should be extended to obtain at least a 16-fold growth in control cultures. The other criteria recommended by the OECD for the algal tests to be acceptable are that the increase in *p*H between the beginning and end of the test should not exceed 1.5 units (for compounds that partly ionize at a *p*H around the test *p*H, the drift should ideally be within 0.5 units); and the coefficient of variation should be ≤10% among the controls. The acceptability of the tests was assessed based on these criteria.

Apart from the test acceptability criteria indicated above, the repeatability of tests was also assessed based on the results obtained from two different experiments using 3,5-dichlorophenol (3,5-DCP) as the reference toxicant. This compound is recommended to be tested at least twice a year to ensure the viability of algal cells by the OECD (2006).

3.2.2.1. Measurement of algal growth. The growth response of *D. tertiolecta* and *C. vulgaris* exposed to each of the studied chemicals was determined by daily measurements of optical density at 680 nm (OD_{680}) with a spectrophotometer (Schimadzu, UV-1208 for experiments conducted with marine algae and Lasany, UV/VIS for experiments conducted

with freshwater algae) over 96 h. The scanning of various wavelengths in the spectrophotometer showed that 680 nm corresponded to the maximum chlorophyll *a* absorption for *D. tertiolecta* as also reported by Janssen et al. (2001). The scanning of various wavelengths in the spectrophotometer also indicated that 680 nm also corresponded to the maximum chlorophyll *a* absorption for *C. vulgaris*; therefore, this wavelength was used to quantify *C. vulgaris* biomass as well.

A linear relationship between algal cell counts and optical density for each species was observed. Therefore, optical density was used as a surrogate measure for the calculation of response variables for *D. tertiolecta* and *C. vulgaris* to express biomass increase during the test. The conversions from optical density to cell counts were done using the linear relationships for specific growth rate calculations. The cell counts were performed using 1 mL of cell suspension which was counted on a gridded cell counter (Sedgewick rafter cell, Sedgewick Co., USA) using Olympus CX41 research microscope (Olympus, Japan). The rafter cell that was used for counting algae holds 100 mm^3 of liquid 1 mm deep over an area of 50x20 mm. The base is divided into 1 mm squares. A cover glass traps the liquid to correct depth. For the determination of the number of algal cells, 5 grids were counted and average of the counts was recorded. The plots of the linear relationship between the optical density at 680 nm (OD_{680}) and the cell counts for *D. tertiolecta* and *C. vulgaris* are provided in Appendix A.

The growth response of *D. tertiolecta* in the presence of phenol and seven chlorophenols (monochlorophenols and dichlorophenols) was also determined by daily measurements of *in vivo* chlorophyll fluorescence intensity based on the visible light emitted by chlorophyll *a* after excitation at 430 nm and emission at 663 nm with a spectrofluorimeter (Perkin-Elmer LS 55). It was found that IC_{50} values obtained either by spectrophotometry or spectrofluorimetry did not differ significantly ($p>0.05$) and their 95% confidence intervals overlapped. Therefore, the IC_{50} and IC_{20} values using OD_{680} values were reported. Geis et al. (2000) also indicated that fluorescence and spectrophotometric wavelengths produced similar variability and inhibition concentrations did not vary significantly. Therefore, we determined the growth of algae based only on spectrophotometric readings throughout the thesis.

3.2.2.2. Analysis of dose-response relationships. The statistic analysis of results was done as suggested by USEPA (2002b). Normality (Shapiro–Wilk's test) and homogeneity of variance (Bartlett's test) were formally tested, since they are the underlying assumptions of the Dunnett's procedure. The lowest observable effect concentration (LOEC) and no observed effect concentration (NOEC) values for growth were obtained using this hypothesis test approach. The NOEC and LOEC of each compound were calculated using Dunnett's test in ToxCalc 5.0.32 (© Tidepool Scientific Software, CA, USA).

To obtain a concentration-response relationship, percent inhibition relative to the control against test substance concentration was fitted using polynomial regression (P.R.) to calculate the IC_{50} and IC_{20} with associated confidence intervals. Curve fitting was carried out in SPSS statistical analysis software (ver. 20.0.0, SPSS, Inc., Chicago, IL, USA) and the roots of the polynomials were calculated using the Scientific Workplace software 3.0 (MacKichan software, WA, USA). Apart from polynomial regression, IC_{50} and associated confidence intervals were also calculated using Weibull and linear interpolation combined with bootstrapping (ICp) methods, as executed in ToxCalc software, to investigate if the toxicity calculation model had a significant impact on the toxicity data. The response variables, the yield and the average specific growth rate, were calculated as recommended in the OECD guideline (2006).

3.3. Modeling Methods

3.3.1. General procedure for QSAR/QSTR modeling

The general procedure for QSAR/QSTR modeling adopted in this thesis is illustrated in Figure 3.3. The modeling process consisted of the following consecutive steps: dataset preparation, division of the dataset into training and test sets, descriptor calculation, modeling on the training set, and eventually testing the predictive performance of the QSAR/QSTR model (Figure 3.3). The details of the software used throughout this thesis can be found in Appendix B.

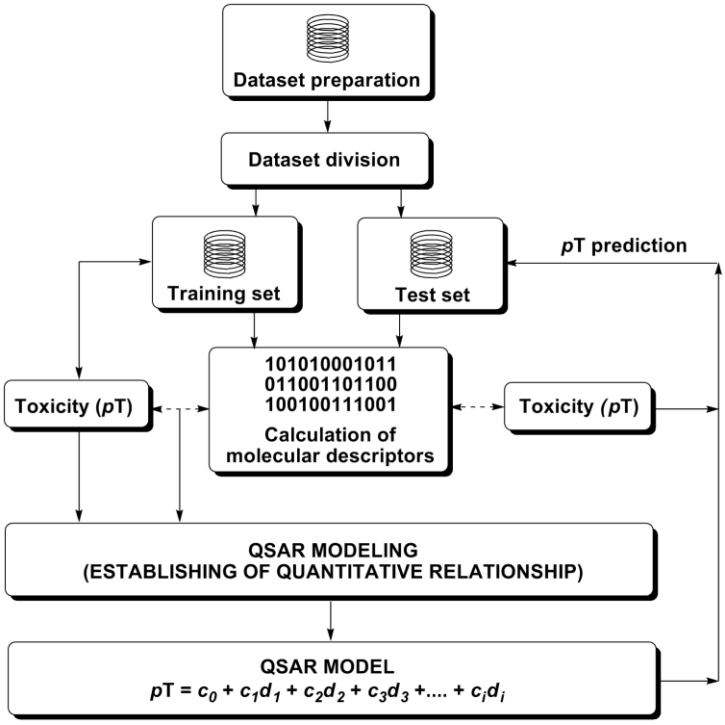


Figure 3.3. The complete QSAR/QSTR methodology workflow (from Saçan et al., 2013)

3.3.2. Molecular descriptors and QSTR development

As a first task, the compounds used in this thesis were sketched in Spartan software (ver. 06, 2006, Wavefunction, Inc., CA, USA) to be able to calculate the molecular descriptors associated with each chemical. Then, geometry optimization of the molecules was carried out in Spartan using semi-empirical PM3 method to determine the lowest energy conformer (hence the most thermodynamically stable) of the phenolic compound in aqueous solution. Quantum chemical descriptors including the energy of the lowest unoccupied molecular orbital (E_{LUMO}), the energy of the highest occupied molecular orbital (E_{HOMO}) were calculated using Spartan software on the energy minimized structures.

To develop global QSTRs, the whole data set was split into training and test sets. The chemicals in the test set were used only for validation purposes and not for model development. This enabled the simulation of a scenario for external validation. To split the compounds into training and test sets, the chemicals were ordered according to their descending toxicity values. Out of 30 compounds, 6 clusters were formed which consisted of 5 compounds. The 3rd compound in each cluster was then assigned to the test set. As a

result, training and test sets were formed with 24 and 6 compounds, corresponding to 80% and 20% of the dataset, respectively. Roy et al. (2008) demonstrated that the external prediction of QSTR models decreased when the ratio of the training set to the whole data set decreased from 80% (75-85 percent in the original work) to smaller ratios. Therefore, it is believed that splitting compounds into the training and test sets with a ratio of 80/20, respectively, was appropriate for QSTR construction.

The logarithms of the octanol/water partition coefficient (Log *P*) as well as the Henry's constant values (H) were extracted from EPI Suite (USEPA, 2011). The negative logarithms of the ionization constant (*p*Ka) values were obtained from OECD (2012). For both Log P and Log *D*, the measured values were used when available.

The distribution coefficient (Log *D*) of the chemicals was calculated at the particular test *p*H using one of the equations provided below:

$$\text{Log } D = \text{Log } P - \log (1+ 10^{pH-pKa}), \text{ for acids (e.g., phenols)} \quad (3.1)$$

$$\text{Log } D = \text{Log } P - \log (1+ 10^{pKa-pH}), \text{ for bases (e.g., anilines)} \quad (3.2)$$

For the investigation of reactivity in the data sets, the algal toxicity data was correlated to the Activation Energy Index (AEI) values of the reactive chemicals. The AEI quantifies the physicochemical properties of the electrophilic oxidation products of *ortho* and *para* substituted polyphenols by describing the difference in combined HOMO and HOMO-1 energies between the electrophile and the intermediate, thus representing the activation energy for reaction. The AEI values were retrieved from Aptula et al. (2005b).

The other molecular descriptors were obtained from Dragon (version 5.4., 2006, Talete, Milano, Italy) and CODESSA Pro (2002, ©University of Florida, Gainesville, USA) software on the geometry optimized structures. 1664 and 818 descriptors were originally retrieved from the Dragon 5.4 and CODESSA Pro software, respectively. Constant descriptors were excluded to minimize the redundant information and the reduced set of 1762 molecular descriptors were imported to CODESSA (ver. 2.2, 1996, ©University of Florida, Gainesville, USA) using an in-house developed software

application, provided kindly by Nikola Minovski from Slovenian Chemical Institute (Minovski, 2009).

All descriptors were analyzed using the CODESSA software with heuristic option, which is a suitable option for descriptor selection (Katritzky et al., 1995a). The heuristic search was restricted to three variables per model to be able to develop simple and informative linear QSTRs. The obtained models at the end of heuristic method were then investigated for a possible co-linearity among the descriptors as judged by the Variance Inflation Factor (VIF) in SPSS. The acceptable MLR models were then used to form a consensus MLR model by averaging the predicted toxicity for every compound.

CP ANN modeling was carried out using the modules developed at the Slovenian National Institute of Chemistry (Zupan and Gasteiger, 1999). For the construction of CP ANN models, the input specifications for a network with 5x5 architecture are summarized in Table 3.5. Note that the number of neurons in x-direction, number of neurons in y-direction and the furthest neuron for correction values change based on the size of the network, e.g., for a 6x6 network, these inputs would take a value of six.

Table 3.5. Details of a sample CP ANN input for a 5x5 network

Property	Value
Number of neurons in x-direction	5
Number of neurons in y-direction	5
Number of weights in each neuron	4
Furthest neuron for correction	5
Minimal correction factor	0.01
Maximal correction factor	0.50
Toroid boundary conditions	No
Type of neighbourhood correction	Triangular
Type of the best match	Neuron with weight most similar to the input
Number of training epochs	100, 200, 400, 800, 1000

The normality of the negative log transformed toxicity data was evaluated using the Kolmogorov-Simirnov test in SPSS. The quality of the models was assessed by the coefficient of determination (r^2), cross-validated regression coefficient (r^2_{cv}) and the root mean square error for the training set ($RMSE_{tr}$). A model with a low cross-validated regression coefficient (r^2_{cv}<0.70), which was used as the criterion for internal validation,

was eliminated. After the elimination of unacceptable models, the predictive ability of the remaining models using Q^2_{F3} formula provided by Consonni et al. (2009):

$$Q^2_{F3} = 1 - \frac{\left[\sum_{i=1}^{n_{test}} \left(y_i - \hat{y}_i\right)^2\right] / n_{test}}{\left[\sum_{i=1}^{n_{training}} \left(y_i - \mu_{tr}\right)^2\right] / n_{training}} \tag{3.3}$$

where y_i and $\hat{y}_i$ are, the observed and predicted (over the test/external set) values of the dependent variable, respectively; μ_{tr} is the mean value of the dependent variable for the training set. A model was considered acceptable if the $Q^2_{F3} > 0.70$.

The external predictive ability of the models were also evaluated on the test set/external set compounds according to the criteria put forward by Golbraikh et al. (2003):

$$R^2 > 0.60 \tag{3.4}$$

$$0.85 \leq k \leq 1.15 \tag{3.5}$$

or

$$0.85 \leq k' \leq 1.15 \tag{3.6}$$

$$\left|R_0^2 - R_{0'}^2\right| \leq 0.3 \tag{3.7}$$

where R^2 is the squared correlation coefficient between observed and predicted values for the test set; R_0^2 and k are the correlation coefficient and slopes of the linear regression between the predicted and observed values, respectively, when intercept was set to zero. The predicted versus observed and observed versus predicted correlation coefficients and slopes are different. Consequently, the latter were designated as $R_0'^2$ and k', respectively.

Finally, the QSTRs were subjected to *Y*-randomization as executed in Molegro Data Modeler (version 2.6.0, 2011, Molegro ApS, Aarhus, Denmark) to ensure that the developed relationships are not chance correlations. In this technique, only the dependent variable is randomly re-ordered while the independent variables are left untouched and a new fit (r_r^2) is obtained for the distorted relationship. In this study, this procedure was

repeated 30 times for each proposed QSTR and the mean r_r^2 was reported for each model. In this so-called model randomization, the resulting models are expected to have lower squared correlation coefficient (r_r^2) compared to the original relationship (r^2) since the link between the structure and activity is severed (Roy et al., 2009). For the evaluation of the randomization, we have used R_p^2 in this study as proposed by Roy et al. (2009) which is calculated by the following formula:

$$R_p^2 = r^2 * (r^2 - r_r^2)^{1/2} \quad (3.8)$$

The R_p^2 should be greater than 0.5 to ensure that model was not obtained by chance.

3.3.3. External validation of the QSTRs

To further test the external predictive ability of the algal QSTRs developed in this study, an external validation set, comprising the 72 h growth inhibition data (pEC$_{50}$, in mmol/L) of 58 compounds towards freshwater alga *P. subcapitata*, was retrieved from literature (Aruoja et al. 2011). The external data set was generated in the same laboratory, by the same researcher following standardized algal growth inhibition test protocols; therefore, it should provide a suitable basis upon which to evaluate the external predictivity of the algal QSARs constructed in this study.

The external data set retrieved from Aruoja et al. (2011) comprises the algal toxicity of 30 phenolic compounds including phenol, chlorophenols, methylphenols and ethylphenols as well as 28 anilines including aniline and its chloro, methyl and ethyl substituted derivatives. The presence of substituents other than chlorine (i.e., methyl and ethyl substituted phenols) and another class of aromatic compounds (i.e., anilines) makes this external validation set highly heterogeneous which, in turn, is believed to be offer a suitable basis to test the predictive ability of the QSTRs.

The performance of the proposed QSTRs in predicting the toxicity of chemicals in the external validation set obtained from Aruoja et al. (2011) was assessed by Q^2_{F3} (Eq. 3.3), Eqs. (3.4-3.7) and the root mean square error of the external validation set ($RMSE_{val}$). Zhu et al. (2008) stated that only a certain fraction of compounds in any external data set is

expected to fall within a model's AD and this fraction is referred to as the data set coverage. Therefore, the proposed QSTRs were also evaluated in terms of their coverage of the compounds in the external data set (% coverage).

3.3.4. Applicability domain

In linear regression analysis it is often required to determine the influence of a given y_i value over each predicted $\hat{y}_i$ value. Contrary to the relationship between the dependent variable and a single predictor (e.g., a molecular descriptor), the influence of multiple predictors on the model might be difficult to determine. Hat matrix, which maps the vector of observed values to the vector of fitted values, can be used to overcome this problem. It is an m × m symmetric matrix where the diagonal elements (known as hat or leverage values) directly reflect the structural influence of a compound to the values predicted by the model (Hoaglin and Welsch, 1978). One of the recommended hat-based methods for applicability domain investigation in case of linear QSTR models is the widely known leverage approach which offers a graphical assessment of the hat values, as a function of the standardized residuals (i.e., the Williams plot) and it is suitable not only for detection of the structurally-influential outliers, but also for determination of the response outliers (Gramatica et al., 2007).

In this study, the leverage approach was used to visualize the applicability domain (AD) of the QSTRs. Compounds with standardized residuals greater than 2.5 standard deviation units were identified as response outliers. The limit of structural outliers was determined by their critical hat values (h*) calculated by $3\,p'/n$, where p' is the number of model variables plus one, and n is the number of compounds in the model. In this approach, the hat value of a particular compound is used as a measure to quantify the compound's distance from the structural space of a model and $h > h^*$ indicate that the compound in question is outside of the model's structural AD; thus, the prediction could be unreliable (Gramatica et al., 2007). The standardized residuals and hat values were calculated in SPSS.

3.3.5. Inter-algal QSTRs

QSTR analysis was conducted in an effort to predict the toxicity of phenols to marine alga *Dunaliella tertiolecta* using the toxicity data of the same set of phenols to freshwater alga *Chlorella vulgaris*. The dataset was also analyzed based on the MOA classification of phenols. Apart from a one-to-one toxicity correlation between *D. tertiolecta* and *C. vulgaris*, molecular descriptors were also used to improve the correlations throughout the analysis.

3.3.6. Interspecies toxicity correlations

The interspecies toxicity correlations were investigated using linear regression analysis. Decimal logarithm of the reciprocal EC_{50} values in mmol/L ($\log(IC_{50})^{-1}$) was used in regression analysis to construct interspecies toxicity correlations. Information regarding the use of toxicity data is provided in Table 3.6.

Table 3.6. The abbreviations used in the interspecies toxicity correlations

Organism	Model variable	Endpoint
Dunaliella tertiolecta	*p*T (Dt)	48 h growth inhibition
Chlorella vulgaris	*p*T (Cv)	96 h growth inhibition
Pseudokirchneriella subcapitata	*p*T (Ps)	72 h growth inhibition
Vibrio fischeri	*p*T (Vf)	30 min. bioluminescence
Tetrahymena pyriformis	*p*T (Tp)	40 h growth inhibition
Daphnia magna	*p*T (Dm)	48 h immobilization
Pimephales promelas	*p*T (Pp)	96 h lethality

The literature toxicity data were retrieved from the following references: 72 h toxicity data of *Pseudokirchneriella subcapitata* was collected from Aruoja et al. (2011). Only the 72 h toxicity of hydroquinone to *P. subcapitata* was obtained from Devillers et al. (1990). 30 min. toxicity data on *Vibrio fischeri* was taken from the TerraTox™ (2006); 40 h toxicity data on *Tetrahymena pyriformis* was collected from Cronin et al. (2002) with the exception of data on 2,3,4,6-tetrachlorophenol and 3,5-dichlorocatechol which were retrieved from Zhu et al. (2008) and Bajot et al. (2011), respectively. The 48 h toxicity data on *Daphnia magna* was retrieved from TerraTox™ (2006) and 96 h toxicity data on *Pimephales promelas* (pT_{Pp}) was collected from Papa et al. (2005). Note that the algal toxicity data, both generated and retrieved from literature, have different endpoint

durations. However, as stated by Eberius et al. (2002), the effect of test duration is eliminated when growth rate is used as the response variable. Therefore, the inter-algal toxicity comparisons and the interspecies toxicity correlations between algae were constructed using data based on specific growth rate calculations.

3.3.7. ECOSAR

From a risk assessment perspective, the widely used ECOSAR (ECOlogical Structure-Activity Relationship Model**)** computer programme for QSAR prediction of chemical toxicity towards aquatic organisms was evaluated by using the novel data sets generated in this thesis. The prediction results of the ECOSAR programme were assessed to evaluate the predictivity and the capability of ECOSAR to correctly classify the chemicals into defined classes of aquatic toxicity according to rules of EU regulation.

The relationships were evaluated by the number of compounds put into analysis (n), square of correlation coefficient (r^2), the leave-one-out cross-validation (r^2_{cv}), and the root mean square error of the model (*RMSE*). An observation was classified as an outlier if the value of its standardized residual was > 2.5.

4. RESULTS AND DISCUSSION

4.1. Toxicity of selected phenols to marine alga *Dunaliella tertiolecta*

All bioassays conducted with marine alga *D. tertiolecta* concurred with the acceptability criteria recommended by the OECD (2006). At the end of 72 h, the mean specific growth rate of algae was found to be 1.07 (± 0.11) d^{-1} which is higher than the minimum specific growth rate (i.e., 0.92 d^{-1}) recommended by the OECD for this exposure duration (OECD, 2006). For all test durations, the coefficient of variation within the controls was ≤10% throughout the tests. The starting *p*H of the bioassays was 8.4 (± 0.1). The *p*H values recorded in the controls at the end of 48 h, 72 h, and 96 h were 8.9 (± 0.1), 9.6 (± 0.1) and 9.9 (± 0.1), respectively.

The GC analysis revealed that only the tested concentration (100 mg/L) of 2-chlorophenol decreased by 5% at the end of 96 h. As stated by Riedl and Altenburger (2007), volatile substances are characterized by a Henry's constant of H ≥ 1 Pa m^3/mol. Based on this criterion, of the chemicals studied, only 2-chlorophenol can be classified as a volatile substance with a Henry's constant of 1.13 Pa m^3/mol. The Henry's constant values of the chlorophenols were retrieved from EPISuite software (The Estimation Programs Interface (EPI) Suite, version 3.20, United States Environmental Protection Agency). It seems that the decrease in the concentration of 2-chlorophenol during the test is due to the volatility of this compound. It should be noted that since none of the test concentrations changed by more than 20%, the nominal concentrations were used to calculate the toxic potency of each compound as recommended by the OECD (2006). The sample chromatogram of 2-chlorophenol as well as the calibration curve used to assess the nominal and measured concentrations of 2-chlorophenol are provided in Appendix C.

The experimental IC_{50} and IC_{20} values and the associated confidence intervals for 48 h, 72 h and 96 h exposures based upon specific growth rate (SGR) and yield are presented in Table 4.1 together with the toxic class, NOEC and LOEC of each test chemical.

Table 4.1. 50% and 20% inhibitory concentrations (IC_{50} and IC_{20}) calculated at the end of 48, 72 and 96 h based on three different methods using yield and specific growth rate (SGR) calculations, no-observed-effect concentration (NOEC), lowest-observed-effect concentration (LOEC), toxic class of phenols for *D. tertiolecta*

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class[b]	NOEC-LOEC[c] (mg/L)
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
P	Yield	ICp	195.4 (126.9-216.3)[a]	105.9 (78.6-169.1)	199.1 (170.4-214.2)	104.2 (66.5-165.0)	199.0 (167.4-214.7)	103.7 (61.4-176.8)	Not toxic	60/120
		Weibull	193.3 (175.7-210.4)	160.5 (124.8-176.4)	191.9 (160.8-214.8)	147.5 (88.0-171.1)	189.1 (146.8-218.7)	134.2 (62.9-163.8)		
		Polynomial Regression	165.4 (148.4-182.2)[c]	83.1 (67.6-103.7)	180.0 (161.2-197.4)	92.1 (67.7-119.3)	187.2 (161.2-212.0)	102.5 (67.3-142.1)		
	SGR	ICp	206.4 (188.8-220.7)	182.5 (95.0-196.8)	222.2 (209.9-232.2)	188.6 (173.7-197.8)	244.2 (236.6-250.0)	193.7 (177.1-205.9)	Not toxic	120/180
		Weibull	208.9 (184.8-196.7)	184.8 (165.3-196.7)	225.8 (194.3-242.5)	192.1 (110.4-211.7)	244.9 (209.8-271.3)	194.0 (120.5-221.1)		
		Polynomial Regression	203.7 (187.2-220.1)	153.4 (133.4-179.7)	217.6 (210.0-224.0)	183.6 (176.4-188.8)	244.2 (236.8-250.6)	193.7 (177.4-206.6)		
2-CP	Yield	ICp	48.8 (35.1-54.1)	13.6 (10.2-18.3)	55.8 (50.7-59.7)	40.0 (9.7-48.9)	60.3 (58.1-62.7)	40.3 (23.0-46.8)	Harmful	20/40
		Weibull	44.0 (18.7-59.6)	24.9 (2.6-37.8)	56.2 (40.7-66.7)	37.3 (14.3-47.6)	58.5 (44.6-68.1)	40.8 (18.1-50.3)		
		Polynomial Regression	46.5 (39.5-54.1)	17.0 (13.3-22.6)	58.6 (51.8-64.2)	29.0 (22.2-37.3)	62.4 (60.1-64.7)	32.1 (27.8-36.6)		
	SGR	ICp	74.8 (64.1-86.2)	43.1 (24.6-49.1)	79.3 (78.5-80.0)	53.2 (46.3-57.4)	78.0 (77.2-78.8)	60.5 (59.1-61.7)	Harmful	20/40
		Weibull	68.4 (50.0-81.2)	46.0 (16.5-58.0)	75.8 (64.2-84.3)	59.2 (35.8-68.0)	77.1 (67.4-84.5)	62.8 (41.0-70.4)		
		Polynomial Regression	71.8 (67.2-76.0)	42.9 (35.4-50.3)	79.5 (75.4-79.7)	57.0 (51.2-59.5)	78.2 (77.0-79.3)	57.6 (56.6-58.6)		
3-CP	Yield	ICp	17.9 (9.3-29.2)	5.5 (3.8-9.2)	30.1 (20.7-44.7)	14.9 (10.1-18.2)	43.6 (35.6-48.8)	25.1 (3.7-38.0)	Harmful	10/20
		Weibull	15.9 (8.0-22.4)	7.2 (1.0-11.5)	29.9 (18.8-39.2)	15.5 (4.5-22.7)	42.1 (34.3-48.3)	30.8 (16.3-36.7)		
		Polynomial Regression	15.9 (14.2-19.1)	5.1 (4.1-6.2)	30.5 (25.0-37.7)	13.3 (10.0-18.5)	41.3 (31.9-49.9)	19.6 (12.4-30.0)		
	SGR	ICp	26.7 (20.2-34.3)	12.8 (6.9-28.9)	47.3 (44.4-50.2)	31.1 (21.9-46.7)	51.2 (46.6-57.4)	42.2 (39.5-44.9)	Harmful	10/20
		Weibull	24.8 (19.3-30.6)	15.3 (8.4-19.6)	46.7 (41.8-51.4)	37.3 (29.4-41.7)	52.3 (44.2-56.5)	43.3 (28.9-48.8)		
		Polynomial Regression	26.5 (23.5-30.2)	11.0 (8.9-13.7)	48.0 (42.5-53.4)	25.4 (20.2-32.2)	54.4 (49.0-61.0)	33.7 (26.7-43.8)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
4-CP	Yield	ICp	16.6 (14.6-18.5)	8.1 (5.1-12.8)	19.9 (17.7-28.9)	14.0 (11.3-14.9)	29.2 (25.9-32.2)	14.2 (11.4-16.3)	Harmful	10/20
		Weibull	17.8 (4.6-28.4)	5.7 (1.0-12.1)	26.7 (15.6-36.4)	12.6 (2.9-19.5)	29.5 (16.9-40.2)	14.0 (3.0-21.8)		
		Polynomial Regression	20.1 (17.8-22.5)	6.4 (5.2-7.8)	27.4 (23.8-32.2)	11.6 (9.2-14.9)	30.4 (28.0-33.0)	12.4 (10.7-14.3)		
	SGR	ICp	49.7 (37.7-75.4)	14.8 (12.5-16.9)	57.1 (52.7-63.3)	19.6 (16.8-24.4)	60.0 (57.6-61.7)	29.8 (26.8-32.7)	Harmful	10/20
		Weibull	41.0 (19.1-59.5)	17.7 (1.1-29.3)	50.4 (33.5-62.0)	31.0 (8.3-41.4)	55.5 (39.2-66.9)	36.1 (11.2-46.5)		
		Polynomial Regression	44.6 (38.0-54.9)	17.1 (14.0-22.1)	54.4 (51.8-57.1)	29.6 (24.9-34.6)	57.0 (56.3-59.3)	34.5 (32.7-37.1)		
2,3-DCP	Yield	ICp	17.1 (13.8-19.4)	10.1 (3.7-18.1)	19.5 (17.6-21.2)	14.3 (8.7-17.3)	22.2 (19.6-24.7)	16.4 (10.9-20.9)	Harmful	10/15
		Weibull	16.7 (12.1-19.8)	11.2 (4.2-14.2)	20.0 (15.8-22.9)	14.5 (7.1-17.4)	22.3 (19.1-24.7)	17.7 (11.0-20.2)		
		Polynomial Regression	17.2 (14.4-19.8)	8.4 (5.9-12.2)	19.8 (19.2-20.4)	11.4 (10.4-12.4)	21.5 (21.1-21.8)	14.1 (13.3-14.8)		
	SGR	ICp	23.0 (22.2-23.9)	16.0 (12.0-18.5)	26.0 (25.7-26.3)	19.3 (17.4-20.9)	26.5 (26.2-26.8)	22.2 (20.0-24.3)	Harmful	10/15
		Weibull	22.5 (18.3-25.5)	16.8 (8.8-19.8)	25.4 (23.4-27.2)	22.0 (17.4-23.8)	26.3 (25.4-28.1)	24.1 (22.0-25.0)		
		Polynomial Regression	22.7 (22.0-23.5)	16.1 (14.2-17.8)	24.7 (24.3-25.2)	19.3 (18.3-20.3)	25.5 (24.9-26.1)	20.8 (19.5-21.9)		
2,4-DCP	Yield	ICp	10.3 (8.1-12.3)	5.5 (3.3-7.2)	11.0 (9.0-12.1)	6.3 (5.3-7.2)	12.1 (9.8-13.1)	7.4 (6.2-9.1)	Harmful	<5/5
		Weibull	10.1 (6.3-12.6)	5.9 (1.5-8.2)	10.7 (7.5-12.9)	6.9 (2.4-9.0)	11.9 (8.7-14.1)	8.1 (3.3-10.2)		
		Polynomial Regression	9.7 (8.4-11.3)	4.1 (3.4-5.2)	10.2 (9.3-11.1)	4.5 (4.0-5.1)	11.7 (10.8-12.7)	5.5 (4.9-6.3)		
	SGR	ICp	15.4 (13.8-16.6)	9.5 (7.3-11.7)	15.8 (15.3-16.3)	11.1 (9.8-11.9)	16.5 (16.2-16.7)	12.5 (11.4-13.5)	Harmful	5/10
		Weibull	14.6 (12.1-16.5)	11.1 (6.3-13.0)	15.4 (13.8-17.1)	12.7 (8.5-14.1)	16.2 (15.3-18.6)	14.1 (11.7-15.0)		
		Polynomial Regression	14.8 (13.5-15.9)	8.0 (6.7-9.5)	15.7 (15.3-16.1)	9.2 (8.7-9.9)	16.7 (16.4-16.9)	10.7 (10.2-11.2)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
2,5-DCP	Yield	ICp	17.0 (16.4-17.3)	14.8 (2.8-15.8)	17.0 (16.5-17.4)	14.1 (10.4-16.1)	17.6 (17.2-18.0)	15.3 (14.9-15.7)	Harmful	10/15
		Weibull	16.9 (n.a)	14.9 (n.a)	17.0 (15.8-18.0)	14.6 (12.7-15.7)	17.8 (16.1-18.9)	15.4 (12.3-16.8)		
		Polynomial Regression	17.0 (15.8-18.0)	11.0 (8.7-13.0)	17.5 (16.3-18.5)	11.9 (9.5-13.9)	18.2 (17.6-18.7)	12.8 (11.7-13.8)		
	SGR	ICp	17.4 (17.3-17.6)	15.7 (15.4-15.9)	18.3 (17.7-19.1)	16.0 (15.7-16.2)	20.5 (19.3-21.3)	17.0 (16.5-17.6)	Harmful	10/15
		Weibull	17.8 (16.8-18.6)	16.2 (14.7-17.1)	18.8 (13.6-19.7)	16.7 (6.4-18.2)	20.4 (n.a)	18.1 (n.a)		
		Polynomial Regression	17.9 (17.6-18.3)	12.9 (12.1-13.7)	19.2 (18.7-19.6)	14.6 (13.7-15.4)	20.3 (20.1-20.6)	16.1 (15.7-16.6)		
2,6-DCP	Yield	ICp	54.1 (48.4-57.9)	33.1 (9.7-52.5)	81.8 (62.7-98.5)	41.9 (19.7-52.9)	97.4 (80.0-104.7)	50.7 (41.3-51.9)	Harmful	<40/40
		Weibull	59.5 (18.7-78.1)	30.8 (1.0-48.1)	77.7 (38.1-104.0)	42.7 (1.5-61.6)	88.3 (58.7-115.5)	53.9 (7.0-71.6)		
		Polynomial Regression	60.9 (54.0-68.6)	24.5 (19.9-30.5)	82.0 (72.9-89.1)	39.2 (30.0-50.1)	91.7 (85.6-97.4)	53.0 (42.6-63.5)		
	SGR	ICp	105.7 (100.3-109.7)	50.2 (43.1-54.7)	116.4 (111.7-123.0)	81.2 (63.2-96.0)	130.1 (117.4-141.5)	100.5 (92.4-102.9)	Not toxic	40/60
		Weibull	99.2 (69.7-202.1)	56.9 (1.0-76.4)	118.1 (102.8-155.5)	89.6 (10.1-96.4)	127.9 (113.7-162.8)	93.6 (67.3-106.1)		
		Polynomial Regression	88.6 (78.4-101.7)	48.7 (32.2-55.5)	110.8 (97.7-128.7)	73.4 (65.3-87.9)	126.9 (110.9-141.3)	86.1 (73.0-105.0)		
3,4-DCP	Yield	ICp	5.4 (4.7-5.9)	4.2 (1.8-5.0)	7.9 (5.6-8.9)	4.4 (1.0-5.6)	8.1 (5.2-8.8)	4.3 (1.0-5.3)	Toxic	4/6
		Weibull	5.7 (4.3-6.6)	3.9 (1.7-4.9)	7.0 (5.3-8.3)	4.8 (1.9-6.0)	7.0 (5.2-8.3)	4.8 (1.8-6.0)		
		Polynomial Regression	6.0 (4.7-7.0)	3.3 (1.9-4.9)	7.3 (6.3-8.0)	4.7 (2.6-6.2)	7.3 (6.6-7.8)	4.8 (3.5-5.9)		
	SGR	ICp	8.3 (8.1-8.4)	5.2 (4.4-5.6)	8.7 (8.5-9.0)	8.0 (5.8-8.4)	8.8 (8.5-8.9)	8.1 (5.8-8.3)	Toxic	4/6
		Weibull	7.7 (6.5-8.6)	6.0 (3.6-6.9)	8.7 (8.3-9.0)	8.0 (7.6-8.3)	8.7 (8.4-9.0)	8.1 (7.8-8.4)		
		Polynomial Regression	7.8 (7.5-8.0)	5.6 (4.9-6.2)	8.4 (8.1-8.6)	6.6 (6.0-7.0)	8.4 (8.2-8.6)	6.7 (6.3-7.0)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
3,5-DCP	Yield	ICp	3.6 (3.2-3.8)	2.5 (1.9-2.8)	6.2 (5.2-6.7)	3.1 (2.4-3.6)	6.8 (6.1-7.3)	4.1 (2.9-5.7)	Toxic	2/4
		Weibull	3.8 (2.6-4.6)	2.4 (0.8-3.2)	5.5 (3.9-6.6)	3.6 (1.2-4.6)	6.7 (5.3-7.6)	4.9 (2.5-5.8)		
		Polynomial Regression	3.2 (2.7-4.0)	1.2 (1.0-1.6)	5.6 (4.7-6.5)	2.3 (1.6-3.4)	6.7 (5.9-7.3)	3.7 (2.6-4.9)		
	SGR	ICp	6.2 (5.2-6.8)	3.0 (2.7-3.3)	8.0 (7.6-8.2)	6.2 (5.9-6.6)	8.5 (8.3-8.7)	7.0 (6.4-7.6)	Toxic	2/4
		Weibull	5.8 (4.3-6.8)	3.9 (1.6-4.9)	7.8 (7.0-8.6)	6.5 (4.4-7.2)	8.5 (8.1-9.3)	7.6 (6.5-7.9)		
		Polynomial Regression	5.6 (4.6-6.7)	2.6 (1.9-3.7)	8.0 (7.7-8.3)	5.8 (5.3-6.3)	8.3 (8.1-8.4)	6.5 (6.1-6.8)		
2,3,4-TCP	Yield	ICp	2.4 (1.4-3.2)	1.2 (0.8-1.4)	4.4 (3.8-5.0)	1.4 (0.7-1.9)	4.7 (3.7-5.3)	1.4 (0.1-2.1)	Toxic	1/2
		Weibull	2.5 (1.2-3.6)	1.0 (0.1-1.8)	3.6 (1.8-5.1)	1.6 (0.2-2.6)	3.9 (1.9-5.5)	1.8 (0.2-2.9)		
		Polynomial Regression	2.6 (2.3-3.1)	1.0 (0.8-1.2)	4.1 (3.5-4.6)	1.5 (1.1-2.1)	4.4 (3.7-5.1)	1.7 (0.9-2.7)		
	SGR	ICp	5.8 (4.8-6.4)	2.0 (1.5-2.7)	6.5 (6.4-6.6)	4.3 (3.8-4.7)	6.6 (6.5-6.7)	4.8 (4.0-5.3)	Toxic	1/2
		Weibull	4.8 (2.9-6.1)	2.7 (0.5-3.8)	6.3 (5.8-7.0)	5.3 (4.0-5.8)	6.5 (6.0-7.1)	5.7 (4.3-6.6)		
		Polynomial Regression	5.2 (4.8-5.5)	2.6 (2.1-3.2)	6.2 (6.0-6.3)	4.3 (4.0-4.6)	6.3 (6.2-6.4)	4.6 (4.3-4.9)		
2,3,5-TCP	Yield	ICp	3.7 (2.4-4.5)	1.0 (0.4-3.5)	3.6 (1.4-4.6)	1.1 (0.7-2.3)	4.3 (4.1-4.5)	1.3 (1.0-1.9)	Toxic	1/2
		Weibull	3.2 (2.0-4.7)	1.4 (0.3-2.1)	3.3 (2.1-4.7)	1.5 (0.4-2.3)	3.6 (2.4-5.1)	1.7 (0.5-2.5)		
		Polynomial Regression	3.3 (2.8-3.9)	1.2 (0.8-1.7)	3.3 (2.8-3.9)	1.3 (1.0-1.8)	4.0 (3.7-4.2)	1.5 (1.3-1.8)		
	SGR	ICp	5.4 (5.2-5.6)	3.1 (1.6-4.2)	5.4 (5.2-5.6)	3.3 (0.8-4.4)	5.6 (5.5-5.7)	4.2 (4.1-4.3)	Toxic	1/2
		Weibull	5.0 (4.3-6.1)	3.6 (2.7-4.1)	5.1 (4.5-6.2)	3.7 (3.0-4.2)	5.3 (4.8-6.6)	4.1 (3.6-4.6)		
		Polynomial Regression	5.5 (5.2-5.8)	3.2 (2.6-3.8)	5.6 (5.3-5.8)	3.4 (2.8-3.9)	5.9 (5.8-6.0)	3.9 (3.7-4.1)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
2,3,6-TCP	Yield	ICp	36.5 (35.0-37.8)	31.1 (29.6-32.5)	50.9 (33.1-52.8)	35.9 (29.1-44.8)	52.4 (48.9-53.6)	35.8 (32.7-48.4)	Harmful	30/40
		Weibull	37.5 (31.4-39.7)	32.6 (18.0-35.7)	48.5 (42.8-52.8)	39.8 (27.3-44.4)	52.0 (46.6-58.7)	46.0 (40.1-50.7)		
		Polynomial Regression	31.2 (27.3-35.6)	26.1 (21.2-30.1)	50.0 (45.5-54.7)	40.6 (36.2-44.6)	51.7 (47.8-55.4)	41.7 (38.8-44.4)		
	SGR	ICp	42.3 (40.3-44.4)	34.3 (33.0-35.9)	53.9 (53.3-54.3)	50.9 (49.9-51.4)	54.3 (53.6-54.7)	50.9 (49.8-51.5)	Harmful	30/40
		Weibull	42.5 (38.3-45.8)	35.7 (27.0-39.2)	53.8 (48.4-58.9)	50.3 (46.3-54.4)	54.6 (49.8-58.8)	51.9 (47.7-55.4)		
		Polynomial Regression	44.4 (42.9-46.1)	33.1 (30.2-35.8)	51.8 (50.6-52.8)	42.4 (39.6-44.5)	52.7 (51.7-53.4)	43.6 (41.1-45.6)		
2,4,5-TCP	Yield	ICp	1.9 (1.6-2.5)	1.1 (0.7-1.5)	2.5 (2.2-2.9)	1.0 (0.4-1.5)	3.3 (2.8-3.8)	1.7 (1.1-2.4)	Toxic	<1/1
		Weibull	2.0 (1.4-2.6)	1.1 (0.4-1.5)	2.4 (1.3-3.4)	1.0 (0.2-1.7)	3.5 (2.0-4.7)	1.7 (0.3-2.6)		
		Polynomial Regression	2.0 (1.7-2.3)	0.8 (0.7-1.0)	2.5 (2.2-2.9)	0.9 (0.8-1.1)	3.7 (3.1-4.4)	1.6 (1.2-2.2)		
	SGR	ICp	3.0 (2.5-3.6)	1.7 (1.3-2.1)	5.0 (4.2-6.2)	2.4 (2.1-2.7)	6.5 (6.3-6.7)	3.3 (2.8-3.9)	Toxic	1/2
		Weibull	3.0 (2.1-3.7)	1.8 (0.7-2.4)	4.7 (3.0-5.8)	2.8 (0.7-3.8)	6.3 (5.6-6.9)	5.1 (3.6-5.7)		
		Polynomial Regression	3.1 (2.7-3.4)	1.4 (1.2-1.6)	5.2 (4.9-5.5)	2.7 (2.3-3.1)	6.1 (5.9-6.3)	4.2 (3.7-4.6)		
2,4,6-TCP	Yield	ICp	36.5 (30.1-43.8)	21.6 (1.0-25.9)	42.9 (42.1-43.7)	31.8 (24.2-38.4)	44.0 (43.3-44.5)	40.4 (37.3-41.3)	Harmful	20/30
		Weibull	34.0 (25.0-40.3)	22.9 (8.4-28.9)	42.5 (40.3-45.4)	37.6 (32.8-39.8)	43.8 (41.4-46.6)	40.5 (36.9-42.3)		
		Polynomial Regression	35.3 (32.3-37.9)	21.1 (15.3-26.6)	40.8 (39.5-41.9)	31.3 (27.9-34.0)	42.0 (41.2-42.8)	32.9 (30.8-34.7)		
	SGR	ICp	43.2 (42.7-43.8)	32.6 (24.3-38.2)	44.4 (44.2-44.6)	41.0 (40.7-41.4)	44.7 (44.6-44.9)	41.5 (41.3-41.8)	Harmful	20/30
		Weibull	42.9 (40.1-44.3)	39.4 (37.1-41.1)	43.5 (41.3-44.9)	39.5 (37.7-41.0)	43.9 (41.4-45.6)	39.4 (36.9-41.1)		
		Polynomial Regression	41.5 (40.5-42.1)	32.0 (30.0-33.5)	43.2 (42.9-43.7)	35.4 (34.4-36.3)	43.6 (43.3-43.9)	35.9 (35.3-36.5)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
3,4,5-TCP	Yield	ICp	1.2 (0.9-1.5)	0.8 (0.6-1.1)	1.3 (0.9-1.6)	0.7 (0.1-1.5)	1.3 (1.1-1.4)	0.4 (0.3-0.6)	Toxic	0.5/1.0
		Weibull	1.0 (0.8-1.2)	0.9 (0.8-1.0)	1.2 (1.0-1.7)	0.8 (0.5-1.0)	1.1 (0.8-1.5)	0.7 (0.3-0.9)		
		Polynomial Regression	1.4 (1.2-1.5)	0.6 (0.4-0.8)	1.2 (1.1-1.3)	0.4 (0.3-0.5)	1.2 (1.1-1.3)	0.4 (0.3-0.5)		
	SGR	ICp	1.3 (0.9-1.6)	0.7 (0.1-1.5)	1.4 (1.3-1.5)	1.1 (1.0-1.2)	1.4 (1.3-1.5)	1.1 (0.9-1.2)	Toxic	0.5/1.0
		Weibull	1.2 (1.0-1.5)	0.8 (0.6-1.0)	1.4 (1.3-1.5)	1.1 (1.0-1.2)	1.3 (1.2-1.4)	1.1 (1.0-1.2)		
		Polynomial Regression	1.5 (1.4-1.6)	1.0 (0.8-1.2)	1.5 (1.4-1.6)	1.0 (0.9-1.1)	1.5 (1.4-1.6)	1.0 (0.9-1.1)		
2,3,4,5-TeCP	Yield	ICp	0.5 (0.3-0.9)	0.2 (0.1-0.3)	0.5 (0.4-1.7)	0.2 (0.1-0.3)	1.0 (0.5-1.3)	0.2 (0.1-0.3)	Very toxic	<0.5/0.5
		Weibull	0.6 (0.1-0.8)	0.2 (0.1-0.4)	0.6 (0.1-0.9)	0.2 (0.1-0.5)	0.8 (0.2-1.2)	0.3 (0.1-0.6)		
		Polynomial Regression	0.6 (0.5-0.7)	0.2 (0.1-0.3)	0.8 (0.7-0.9)	0.2 (0.1-0.3)	0.9 (0.8-1.0)	0.3 (0.2-0.4)		
	SGR	ICp	1.12 (1.1-1.2)	0.4 (0.2-0.9)	1.5 (1.4-1.6)	0.5 (0.4-1.7)	1.6 (1.5-1.7)	1.1 (0.7-1.3)	Toxic	<0.5/0.5
		Weibull	1.1 (0.8-1.2)	0.8 (0.4-0.9)	1.4 (1.1-1.6)	1.0 (0.5-1.2)	1.6 (1.4-1.8)	1.2 (1.0-1.4)		
		Polynomial Regression	0.94 (0.86-1.03)	0.40 (0.35-0.47)	1.39 (1.28-1.50)	0.65 (0.54-0.78)	1.60 (1.55-1.65)	0.92 (0.84-1.01)		
2,3,4,6-TeCP	Yield	ICp	1.1 (0.8-1.4)	0.4 (0.2-0.9)	2.1 (1.2-2.8)	0.7 (0.4-0.9)	1.9 (1.1-2.6)	0.3 (0.1-0.7)	Toxic	0.5/1.0
		Weibull	1.0 (0.5-1.6)	0.4 (0.1-0.8)	1.8 (0.9-2.9)	0.7 (0.1-1.2)	1.9 (0.8-3.2)	0.6 (0.1-1.2)		
		Polynomial Regression	1.2 (1.0-1.5)	0.4 (0.2-0.5)	2.1 (1.7-2.6)	0.7 (0.5-1.0)	3.3 (2.8-3.8)	1.5 (1.1-2.1)		
	SGR	ICp	2.8 (0.9-3.8)	1.0 (0.7-1.3)	4.4 (4.0-4.7)	2.0 (1.2-2.7)	4.7 (4.4-4.9)	2.6 (0.1-5.6)	Toxic	0.5/1.0
		Weibull	2.5 (1.4-3.6)	1.0 (0.2-1.7)	4.2 (3.5-4.8)	3.1 (1.7-3.7)	4.6 (4.2-5.1)	3.8 (3.5-4.2)		
		Polynomial Regression	2.9 (2.4-3.4)	1.0 (0.8-1.4)	4.2 (3.9-4.4)	2.4 (1.9-2.9)	4.5 (4.1-4.8)	3.0 (2.1-3.7)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
2,3,5,6-TeCP	Yield	ICp	1.0 (0.9-1.1)	0.4 (0.2-0.8)	1.4 (0.5-3.1)	0.3 (0.2-0.5)	2.0 (1.0-4.4)	0.8 (0.1-1.7)	Toxic	<0.5/0.5
		Weibull	1.1 (0.5-1.5)	0.4 (0.1-0.7)	1.2 (0.1-2.3)	0.3 (0.1-0.7)	2.2 (0.7-4.3)	0.7 (0.1-4.3)		
		Polynomial Regression	1.1 (1.0-1.3)	0.4 (0.3-0.5)	1.8 (1.4-2.2)	0.4 (0.2-0.7)	2.9 (2.2-3.4)	1.0 (0.6-1.7)		
	SGR	ICp	2.4 (2.1-2.7)	1.4 (1.2-1.7)	4.2 (3.7-4.7)	2.3 (0.8-4.0)	5.4 (4.9-5.6)	2.8 (0.6-5.4)	Toxic	2/3
		Weibull	2.5 (1.8-2.9)	1.7 (0.7-2.1)	4.1 (3.4-5.8)	2.6 (1.3-3.2)	5.2 (3.8-11.2)	3.7 (1.5-4.9)		
		Polynomial Regression	2.6 (2.4-2.8)	1.6 (1.3-1.8)	4.9 (4.3-5.4)	2.8 (1.8-3.4)	6.0 (5.8-6.1)	4.2 (3.9-4.6)		
PCP	Yield	ICp	0.23 (0.16-0.28)	0.15 (0.11-0.19)	0.27 (0.26-0.28)	0.16 (0.01-0.18)	0.32 (0.30-0.33)	0.19 (0.16-0.22)	Very toxic	0.1/0.2
		Weibull	0.24 (0.17-0.29)	0.16 (0.06-0.20)	0.27 (0.18-0.33)	0.17 (0.05-0.23)	0.31 (0.23-0.37)	0.21 (0.08-0.27)		
		Polynomial Regression	0.25 (0.21-0.29)	0.13 (0.09-0.17)	0.29 (0.27-0.31)	0.14 (0.11-0.18)	0.33 (0.32-0.35)	0.19 (0.16-0.22)		
	SGR	ICp	0.35 (0.34-0.37)	0.21 (0.15-0.26)	0.41 (0.40-0.42)	0.27 (0.25-0.28)	0.43 (0.42-0.44)	0.32 (0.31-0.33)	Very toxic	0.1/0.2
		Weibull	0.34 (0.26-0.40)	0.24 (0.11-0.30)	0.40 (0.34-0.44)	0.32 (0.21-0.36)	0.42 (0.40-0.45)	0.37 (0.32-0.39)		
		Polynomial Regression	0.36 (0.34-0.37)	0.23 (0.20-0.26)	0.39 (0.38-0.40)	0.28 (0.27-0.29)	0.41 (0.40-0.42)	0.31 (0.30-0.32)		
Pyrogallol	Yield	ICp	31.6 (26.7-36.2)	14.0 (5.0-20.4)	35.0 (30.2-39.4)	10.1 (4.3-17.1)	38.3 (33.1-42.2)	13.1 (0.8-22.2)	Harmful	4/8
		Weibull	29.4 (21.1-39.6)	15.7 (5.8-21.7)	30.7 (21.7-41.5)	16.0 (5.9-22.4)	33.5 (24.9-43.8)	18.8 (8.2-25.3)		
		Polynomial Regression	32.5 (29.7-35.4)	13.7 (9.8-17.7)	34.3 (31.5-37.0)	14.9 (11.1-18.6)	38.0 (33.1-42.0)	17.1 (11.6-23.5)		
	SGR	ICp	43.2 (41.8-44.2)	28.4 (23.9-32.4)	45.0 (44.1-45.9)	33.6 (32.3-35.0)	45.9 (44.9-46.7)	35.1 (33.5-36.3)	Harmful	4/8
		Weibull	40.5 (35.8-52.2)	30.1 (24.0-34.0)	43.0 (38.4-54.8)	33.8 (30.2-37.7)	44.0 (39.3-58.3)	35.7 (32.7-40.4)		
		Polynomial Regression	45.3 (43.8-46.7)	28.5 (25.2-31.5)	47.3 (46.0-48.4)	32.2 (29.4-34.7)	48.3 (47.5-49.1)	34.2 (32.4-36.0)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
Hq	Yield	ICp	0.32 (0.21-0.35)	0.16 (0.11-0.26)	0.33 (0.24-0.36)	0.17 (0.08-0.28)	0.33 (0.25-0.36)	0.17 (0.13-0.22)	Very toxic	0.1/0.2
		Weibull	0.29 (0.23-0.34)	0.21 (0.11-0.26)	0.31 (0.21-0.37)	0.20 (0.06-0.26)	0.31 (0.21-0.38)	0.19 (0.06-0.26)		
		Polynomial Regression	0.30 (0.22-0.37)	0.17 (0.09-0.28)	0.33 (0.29-0.38)	0.19 (0.11-0.27)	0.34 (0.30-0.36)	0.19 (0.14-0.24)		
	SGR	ICp	0.34 (0.32-0.36)	0.29 (0.08-0.33)	0.42 (0.39-0.45)	0.31 (0.23-0.35)	0.45 (0.44-0.47)	0.32 (0.27-0.35)	Very toxic	0.1/0.2
		Weibull	0.34 (0.32-0.36)	0.3 (0.26-0.32)	0.42 (0.34-0.52)	0.3 (0.1-0.4)	0.46 (0.38-0.69)	0.31 (0.07-0.38)		
		Polynomial Regression	0.34 (0.31-0.39)	0.23 (0.17-0.30)	0.43 (0.41-0.45)	0.29 (0.24-0.34)	0.46 (0.45-0.47)	0.31 (0.28-0.34)		
ClHq	Yield	ICp	0.38 (0.33-0.43)	0.23 (0.14-0.34)	0.40 (0.34-0.42)	0.18 (0.01-0.37)	0.41 (0.39-0.42)	0.25 (0.01-0.39)	Very toxic	0.1/0.2
		Weibull	0.36 (0.29-0.41)	0.26 (0.13-0.32)	0.36 (0.28-0.42)	0.25 (0.11-0.31)	0.39 (0.33-0.43)	0.30 (0.19-0.35)		
		Polynomial Regression	0.37 (0.34-0.40)	0.24 (0.17-0.30)	0.37 (0.34-0.40)	0.23 (0.18-0.29)	0.38 (0.36-0.40)	0.26 (0.21-0.30)		
	SGR	ICp	0.43 (0.42-0.44)	0.35 (0.30-0.41)	0.44 (0.43-0.45)	0.40 (0.38-0.41)	0.44 (0.43-0.45)	0.41 (0.40-0.42)	Very toxic	0.2/0.3
		Weibull	0.43 (0.41-0.46)	0.38 (0.34-0.40)	0.44 (0.41-0.48)	0.40 (0.38-0.41)	0.44 (0.42-0.48)	0.40 (0.38-0.42)		
		Polynomial Regression	0.41 (0.40-0.42)	0.32 (0.30-0.34)	0.42 (0.41-0.43)	0.33 (0.32-0.34)	0.43 (0.42-0.44)	0.34 (0.33-0.35)		
TetraClHq	Yield	ICp	5.1 (4.5-5.6)	2.2 (1.2-3.6)	4.9 (3.2-5.7)	2.6 (1.6-3.8)	4.8 (2.6-5.8)	2.9 (2.0-4.1)	Toxic	2/4
		Weibull	4.6 (3.8-5.7)	3.0 (1.8-3.7)	4.6 (3.9-6.1)	3.1 (1.8-3.7)	4.6 (4.0-6.2)	3.3 (2.0-3.7)		
		Polynomial Regression	5.2 (4.6-5.6)	2.7 (2.0-3.4)	5.4 (5.0-5.8)	3.3 (2.4-4.0)	5.7 (5.4-5.9)	3.6 (3.1-4.1)		
	SGR	ICp	5.7 (5.6-5.9)	4.4 (4.1-4.6)	5.7 (5.4-5.9)	4.4 (3.9-4.7)	5.8 (5.5-6.0)	4.4 (4.0-4.7)	Toxic	2/4
		Weibull	5.5 (4.9-7.3)	4.4 (4.1-5.0)	5.5 (4.8-10.2)	4.5 (4.1-5.7)	5.5 (4.6-5.9)	4.5 (4.1-n.a)		
		Polynomial Regression	6.0 (5.9-6.2)	4.3 (4.0-4.5)	6.1 (5.9-6.3)	4.4 (4.1-4.7)	6.1 (6.0-6.3)	4.5 (4.1-4.8)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
Cat	Yield	ICp	17.6 (15.8-19.4)	11.4 (8.2-14.3)	19.7 (17.9-22.4)	12.9 (0.1-14.5)	23.6 (15.9-30.0)	13.4 (8.0-15.8)	Harmful	10/20
		Weibull	18.4 (3.3-21.9)	13.4 (0.1-16.5)	22.4 (15.6-30.1)	11.6 (4.2-16.4)	26.3 (17.4-37.1)	12.6 (3.5-18.6)		
		Polynomial Regression	20.3 (17.8-23.1)	9.8 (7.9-11.9)	23.3 (21.2-25.7)	11.0 (9.3-12.9)	27.4 (24.7-30.4)	12.2 (10.3-14.5)		
	SGR	ICp	28.7 (23.7-35.4)	15.6 (13.3-17.6)	44.2 (41.3-46.7)	20.5 (18.2-23.1)	51.9 (50.1-53.6)	27.3 (22.4-32.9)	Harmful	10/20
		Weibull	29.0 (20.9-37.0)	16.7 (6.4-22.5)	40.9 (31.6-53.2)	24.4 (11.0-31.6)	48.2 (41.3-59.8)	33.4 (23.0-39.3)		
		Polynomial Regression	29.4 (24.3-36.3)	13.4 (11.1-17.0)	45.7 (43.7-47.5)	21.5 (18.3-24.5)	53.8 (52.3-55.1)	30.6 (27.4-33.5)		
4-ClCat	Yield	ICp	20.6 (17.6-22.6)	12.3 (6.5-16.3)	25.1 (24.0-26.1)	9.8 (7.8-32.5)	28.0 (26.1-29.4)	21.0 (1.0-22.8)	Harmful	5/10
		Weibull	20.0 (15.2-23.4)	13.9 (6.0-17.2)	25.5 (21.4-28.2)	20.0 (12.5-23.1)	28.4 (23.7-31.7)	22.3 (11.7-25.6)		
		Polynomial Regression	19.8 (17.1-20.4)	12.7 (9.3-15.0)	26.2 (23.6-28.2)	15.2 (10.9-19.2)	29.1 (26.5-31.2)	19.6 (14.3-23.9)		
	SGR	ICp	26.9 (25.2-29.3)	20.5 (18.3-21.6)	32.2 (31.7-32.9)	24.7 (23.7-26.0)	33.8 (33.4-34.1)	30.0 (28.0-30.5)	Harmful	5/10
		Weibull	27.1 (22.2-30.4)	20.8 (10.9-24.3)	31.7 (29.9-43.4)	27.5 (2.0-29.3)	33.3 (30.6-35.4)	30.1 (28.8-32.1)		
		Polynomial Regression	28.3 (27.0-29.5)	18.5 (16.2-20.8)	31.8 (31.3-32.3)	24.4 (23.2-25.4)	33.1 (32.7-33.5)	26.5 (25.5-27.4)		
3,5-DiClCat	Yield	ICp	11.6 (9.5-13.5)	6.9 (4.7-7.9)	15.8 (12.9-18.1)	8.6 (5.3-12.8)	19.0 (16.7-22.2)	10.2 (5.1-13.4)	Harmful	5/10
		Weibull	11.7 (8.7-15.0)	6.8 (3.0-9.0)	15.9 (11.5-20.4)	9.3 (3.1-12.5)	18.9 (14.2-25.0)	11.1 (4.1-14.6)		
		Polynomial Regression	11.5 (10.1-13.2)	5.0 (3.9-6.3)	16.6 (13.6-19.6)	7.3 (5.2-9.8)	20.7 (19.1-22.3)	9.3 (7.1-11.5)		
	SGR	ICp	17.7 (14.8-21.5)	10.5 (9.1-11.7)	26.0 (24.5-27.8)	16.0 (13.7-19.0)	27.8 (27.2-28.5)	20.6 (19.6-21.6)	Harmful	5/10
		Weibull	17.7 (13.2-23.0)	10.9 (3.1-14.2)	24.3 (21.3-34.9)	17.7 (12.1-20.2)	26.3 (23.3-40.3)	20.6 (18.1-23.3)		
		Polynomial Regression	19.4 (16.6-22.3)	8.6 (7.0-10.8)	27.7 (26.4-28.8)	16.9 (14.6-19.0)	29.5 (29.0-30.0)	20.2 (19.0-21.3)		

Table 4.1. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
Res	Yield	ICp	286.3 (267.2-305.4)	202.6 (150.4-222.9)	310.3 (291.5-328.1)	179.7 (77.9-251.4)	309.4 (294.1-325.8)	216.5 (36.6-271.1)	Not toxic	100/200
		Weibull	281.9 (230.8-318.3)	212.9 (114.3-250.0)	302.2 (259.4-336.1)	238.1 (147.9-271.6)	307.6 (275.6-341.9)	253.7 (162.2-280.7)		
		Polynomial Regression	334.9 (314.6-333.5)	263.3 (239.3-267.4)	325.1 (327.7-334.4)	261.4 (255.7-276.2)	335.2 (321.0-350.4)	272.6 (251.0-294.2)		
	SGR	ICp	333.5 (328.9-337.9)	273.8 (250.2-293.1)	341.7 (337.9-344.5)	306.8 (300.6-311.2)	342.4 (339.8-344.4)	307.8 (303.7-311.1)	Not toxic	100/200
		Weibull	328.6 (311.8-446.6)	291.9 (243.8-306.5)	337.4 (333.9-341.2)	299.1 (290.3-306.7)	339.4 (335.0-345.4)	312.0 (302.4-316.7)		
		Polynomial Regression	334.9 (314.6-335.5)	263.3 (239.3-267.4)	325.1 (317.7-334.4)	261.4 (245.7-276.2)	335.2 (321.0-350.4)	272.6 (251.0-294.2)		
4-ClRes	Yield	ICp	102.8 (75.1-130.7)	36.1 (3.8-75.5)	122.5 (105.6-134.1)	69.7 (41.5-90.3)	135.0 (124.0-144.5)	92.4 (62.0-115.2)	Not toxic	50/100
		Weibull	92.1 (61.4-124.6)	48.0 (12.5-68.5)	116.2 (100.6-172.0)	81.7 (43.7-95.3)	120.1 (110.4-128.7)	85.8 (78.5-91.1)		
		Polynomial Regression	99.5 (86.8-112.4)	38.4 (21.7-55.3)	145.1 (138.6-148.7)	94.1 (79.5-105.3)	156.4 (154.6-158.3)	118.7 (114.3-122.8)		
	SGR	ICp	137.5 (130.9-144.3)	100.0 (64.5-110.8)	144.0 (141.1-146.4)	110.4 (105.7-114.3)	146.8 (144.6-148.6)	115.0 (111.4-117.7)	Not toxic	50/100
		Weibull	129.7 (115.4-165.2)	98.5 (82.9-110.3)	137.7 (118.7-475.2)	113.3 (104.3-182.8)	140.4 (137.8-143.5)	116.7 (112.2-121.3)		
		Polynomial Regression	145.1 (138.6-148.7)	94.1 (79.5-105.3)	152.4 (149.2-155.7)	111.7 (104.5-118.5)	156.4 (154.6-158.3)	118.7 (114.3-122.8)		
4,6-DiClRes	Yield	ICp	53.1 (45.3-58.7)	15.0 (2.3-26.0)	52.8 (33.8-60.3)	16.8 (9.9-25.5)	54.0 (38.7-60.6)	18.3 (8.8-71.0)	Harmful	40/80
		Weibull	48.6 (42.4-52.2)	30.9 (25.6-34.9)	53.2 (47.8-55.9)	39.2 (36.4-44.3)	53.4 (50.3-57.4)	39.8 (36.6-44.4)		
		Polynomial Regression	54.9 (50.9-60.5)	30.9 (20.4-37.8)	58.5 (54.9-61.7)	38.5 (31.4-44.1)	58.3 (54.1-62.1)	38.1 (28.0-46.2)		
	SGR	ICp	58.0 (55.3-59.5)	44.8 (41.0-47.4)	58.1 (53.6-60.3)	44.9 (37.8-48.5)	58.5 (56.0-60.1)	45.6 (41.5-48.2)	Harmful	40/80
		Weibull	57.3 (54.9-59.4)	47.6 (43.3-50.1)	57.5 (55.0-59.6)	48.1 (44.4-52.2)	58.2 (56.6-60.2)	48.4 (46.2-50.1)		
		Polynomial Regression	60.9 (58.6-62.7)	43.4 (37.9-47.8)	61.5 (59.2-63.4)	44.8 (39.4-49.2)	62.6 (61.8-63.4)	47.3 (45.4-49.0)		

[a]:confidence intervals; [b]: toxic class based on 72h ICp calculations; [c]: NOEC and LOEC values were calculated using 72h toxicity data

D. tertiolecta revealed dose-dependent responses to chemicals tested in this study. The response of algae to phenol is provided as an example in Figure 4.1.

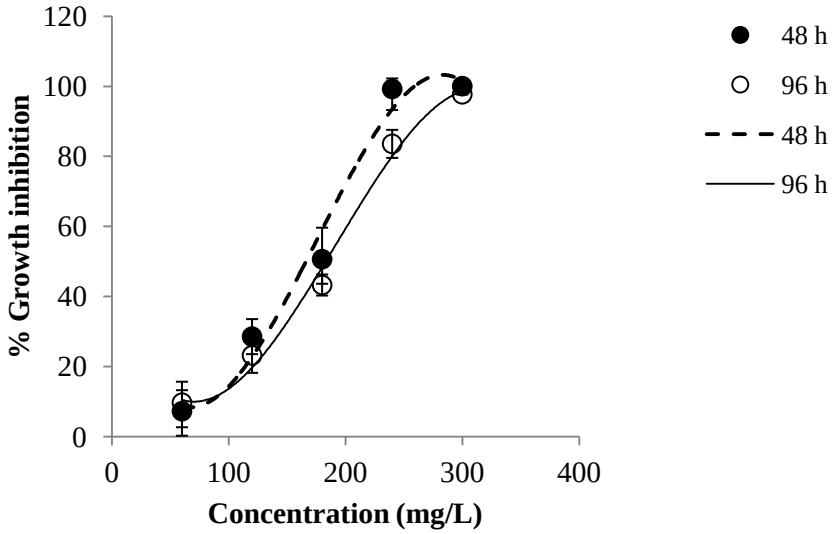


Figure 4.1. Growth response of *Dunaliella tertiolecta* exposed to different concentrations of phenol (bars represent experimental errors)

Based upon average specific growth rate, the IC_{50} and associated confidence intervals for 48 h and 96 h did not overlap, which suggests that the toxicity of phenol and chlorophenols to *D. tertiolecta* tended to decrease between these durations (Table 4.1). This could be attributed to the increase in *p*H of the test medium because of the fixation of CO_2 during photosynthesis. This, in turn, affects the uptake, bio-concentration and toxicity of phenolic compounds (Neuwoehner and Escher, 2001). Weak acids such as chlorophenols tend to ionize at a *p*H greater than their acid dissociation constants (*pK*a) and the degree of ionization enhances as the (*p*H – *pK*a) difference increases. The decrease in toxicity of weak acids has been attributed to the fact that unionized form of the molecule contributes to the toxicity more than the ionized form because the neutral molecule is more bioavailable than the corresponding charged molecule (Fahl et al., 1995; Escher and Schwarzenbach, 2002). For example, Escher and Hermens (2004) pointed out that the toxicity of 2,4,5-trichlorophenol to *Scenedesmus vacuolatus* decreased by 10-fold as the *p*H of the test medium increased from 6.4 to 8.1. Fahl et al. (1995) demonstrated that the toxicity of sulfonylurea herbicides, which are weak acids like chlorophenols, was lower in *p*H 6 than in *p*H 5 to the freshwater alga *Chlorella fusca*. In another study, Lee and Chen (2009) determined the toxicity of benzoic acids to *P. subcapitata* and also found that *p*H

increase led to a reduction in the toxicity of these chemicals. It is, therefore, likely that the *p*H increase caused by algal growth rendered the chlorophenols less toxic to *D. tertiolecta* as exposure duration increased from 48 h to 96 h. Since the influence of *p*H on toxicity would be minimal at the end of 48 h exposure, the toxicity estimates for this exposure duration can be used to denote the toxicity of chlorophenols towards *D. tertiolecta*, if the influence of *p*H on toxicity is desired to be kept at minimum.

Another factor that can be related to the decrease in toxicity of chlorophenols with increasing endpoint duration might be the adaptation/acclimation of algae to the test compounds. Throughout the tests, at higher test concentrations, a lag period was observed before the algal cells resumed growth. Similar observations were also reported by Olivier et al. (2003) who stated that the *Chlorella* VT-1 adapted to chlorophenol compounds and after a lag period the cultures began to grow rapidly. Authors suggested that the lag phase indicates some form of detoxification which is required before the algae can resume growth. In the same context, Scragg et al. (2003) suggested that resistant cells survive and initiate the later growth. Another observation was highlighted by Scragg (2006) that naphthalene inhibited the growth of approximately 98% of *Chlamydomonas angulosa* algal cells on addition, but after 3 d the growth was restored and matched the controls. We also observed later growth of *D. tertiolecta* cells after a lag phase in the presence of relatively high concentrations of each test chemical. In conclusion, together with the influence of *p*H on toxicity as discussed above, algal adaptation to chlorophenols might have a role in rendering these compounds less toxic at the end of 96 h exposure as compared to 48 h exposure.

Based on the IC_{50} values, the least toxic compound was found to be resorcinol, while the most toxic compound was pentachlorophenol regardless of exposure duration or response variable (Table 4.1). The toxicity of the chemicals based on IC_{20} values also followed the same toxicity pattern towards *D. tertiolecta*. As expected, either the IC_{50} or IC_{20} values based upon average specific growth rate were found to be higher than those based upon yield due to the mathematical basis of the respective approaches (OECD, 2006). It was also observed that the toxicity calculation method (i.e., ICp, Weibull or polynomial regression) did not change the IC_{50} significantly as the confidence intervals for the toxicity data overlapped in most of the cases.

The European Union Technical Guidance Document (EU TGD, 2003) classifies the toxicity of chemicals to aquatic organisms according to the EC_{50} values (effective concentration that reduces the measured endpoint by 50% and the endpoint encompasses lethality, immobilization, growth inhibition, etc.). Within this scheme, the toxicity of compounds is classified as outlined in Table 4.2.

Table 4.2. Toxicity classification of chemicals according to EU TGD (2003)

Toxicity range (mg/L)	**Class**
$EC_{50} \leq 1$	Very toxic
$1 < EC_{50} \leq 10$	Toxic
$10 < EC_{50} \leq 100$	Harmful
$EC_{50} > 100$	Not toxic (not classified)

As can be seen from Table 4.1, IC_{50} values calculated by the yield method revealed that 13 chemicals were classified as harmful, nine chemicals as toxic, two chemicals as very toxic and four chemicals was as not toxic according to the toxicity scheme outlined in Table 4.2. When the IC_{50} calculated by the average growth rate method were used for classification, the toxic class of 2,6-DCP and 2,3,4,5-TeCP was reduced from harmful to not classified and from very toxic to toxic, respectively. No matter which response variable was selected for toxicity classification, 3,5-DCP and chlorophenols having at least three chlorine atoms (with the exception of 2,3,6-TCP and 2,4,6-TCP), were found to be either toxic or very toxic to *D. tertiolecta* (Table 4.1). It should be kept in mind that environmental factors (e.g., *p*H, suspended matters, temperature and the presence of other chemicals) may enhance the acute or chronic toxicity of these chemicals; consequently, their release to the environment may cause irreversible adverse effects on algal populations. Moreover, if algal growth is affected, the biomass at higher tropic levels can be impacted as well (Saçan and Balcıoğlu, 2006). On the other hand, although the typical chlorophenol concentrations reported in the aquatic environment are not higher (Czaplicaka, 2004) than the NOEC or IC_{20} values reported in the present study, long-term effects of continuous low-level exposure to chemicals might also have unexpected consequences on the ecosystem (Saçan and Balcıoğlu, 2006).

Regardless of the exposure duration, the results revealed that the toxicity of chlorophenols generally increased with increasing number of chlorine atoms on the aromatic ring, which is consistent with the previous literature reports (USEPA, 1980; Boyd

et al., 2001). Additionally, the –o*rtho* substituted chlorophenols were found to be less toxic than the –*meta* and –*para* substituted congeners. As stated by Boyd et al. (2001), this has been ascribed to the shielding of the OH group by –*ortho* substituted chlorine atom(s). Among mono-chlorophenols, 2-CP was found to be less toxic than 3-CP or 4-CP. The simultaneous occurrence of chlorine atoms substituted at the 2- and 6- positions is known to result in a lower toxicity of these congeners (Briens et al., 1999; Escher and Schwarzenbach, 2002; Czaplicka, 2004). Consistent with the –*ortho* effect mentioned above, 2,6-DCP was the least toxic dichlorophenol tested. Of the trichlorophenols, 2,3,6-TCP and 2,4,6-TCP were found to be less toxic than 2,3,4-TCP or 2,4,5-TCP and for tetrachlorophenols, 2,3,4,6-TeCP and 2,3,5,6-TeCP were both less toxic than 2,3,4,5-TeCP.

The toxicity of 3,5-DCP, the reference toxicant, was determined twice for marine alga *D. tertiolecta* and it was found that the 95% confidence intervals overlapped for the two IC_{50} values. This indicates that the algal response did not change significantly during the course of the study. Additionally, the 72 h IC_{50} of 3,5-DCP determined in the present study (5.6 ± 0.9 mg/L based upon yield and 8.0 ± 0.3 mg/L based upon average specific growth rate) was comparable to the toxicity of this compound (6.4 ± 2.4 mg/L) to the freshwater alga *Desmodesmus subspicatus* determined in a 72 h algal growth inhibition ring test participated by 18 laboratories (ISO, 2004).

4.1.1. Correlation of *D. tertiolecta* toxicity with hydrophobicity

Chlorophenols up to four chlorine substitution as well as resorcinols are expected to exhibit toxicity mainly by polar narcotic mechanisms and consequently the toxicity of these compounds to marine algae should correlate well with hydrophobicity as described by Log *P* (Vighi et al., 2009). Previous findings in the literature also suggest that the toxicity of chlorophenols with respiratory uncoupling ability (i.e., tetra-chlorophenols and pentachlorophenol) may have good correlation to hydrophobicity as well (Pepelko et al., 2005).

Analysis revealed that there is indeed a strong trend between hydrophobicity as described by Log *P* and the marine algal toxicity of polar narcotics and respiratory

uncouplers determined at the end of 48 h (n=23, r^2= 0.85, r^2_{cv}=0.82). However, the toxicity of 2,6-DCP, 2,3,6-TCP and 2,4,6-TCP was over-estimated by the hydrophobicity model and the standardized residuals of these compounds were found to be higher than the cut-off value of 2.5. Removal of 2,6-DCP, 2,3,6-TCP and 2,4,6-TCP from the analysis significantly increased the correlation between the hydrophobicity and marine algal toxicity (Eq. 4.1):

$$pT_{48} = 0.83 \text{ Log } P - 1.44 \quad (4.1)$$

(n=20, r^2= 0.96, r^2_{cv}=0.95)

As discussed previously, chlorine atoms located at the 2- and 6- positions relative to the OH group leads to the lower the toxicity of this compound compared to the congeners with equal number of chlorine atoms attached to the aromatic ring. As explained by Escher et al. (1999), this is probably because heterodimer formation is hindered due to localization of the charge as well as steric interactions by the simultaneous occupation of both *ortho* positions in a chlorophenol molecule which in turn reduces its toxicity. Interestingly, the toxicities of three other compounds used in the present study which have chlorine atoms substituted simultaneously at 2- and 6-positions (2,3,4,6—TeCP, 2,4,5,6-TeCP and PCP) had standardized residuals lower than one (results not shown), indicating that the models predicted the toxicities of these compound well. Czaplicka (2004) stated that the toxicity of chlorophenols increases if chlorine atoms are substituted at the 3-, 4-and 5- positions. Based on this information, it might be possible that the presence of chlorine atoms at 4- and 5- positions in 2,4,5,6-TeCP and at 3-, 4-, and 5- positions in PCP counteracts the lowering effect of 2- and 6- positions in the toxicity of these molecules towards *D. tertiolecta*. Additionally, as stated by Schüürman et al. (1997), the high lipophilicity of PCP (Log *P* of 5.12) leads to high narcotic-type membrane perturbations in the organisms and this might be the reason why it is the most toxic chlorophenol to *D. tertiolecta*. Another reason might be the strong uncoupling activity of tetra-chlorophenols and pentachlorophenol which might have resulted in an increased toxicity as compared to narcosis. It should also be kept in mind that the absence of a cell wall in *D. tertiolecta* might facilitate the transport of pollutants to the cytoplasm; consequently, the proportional increase in the marine algal toxicity of polar narcotic phenols with increasing hydrophobicity can be linked to the physiological characteristics of *D. tertiolecta*.

From a regulatory perspective, EU TGD (2003) proposed several equations employing the hydrophobicity parameter, Log *P*, for the use of these relationships in aquatic risk assessment. Although there is a benchmark equation to be used in risk assessment for freshwater algae (reported toxicity to: *Pseudokirchneriella subcapitata*) regarding non-polar narcotics, there is no equation available for polar narcotics. As for marine algae, there is no equation either for non-polar or polar narcotics. It was underlined in the technical guidance document that only the experimental data that were generated according to OECD test guidelines or comparable methods were used in the reported models. Since the toxicological assays in this study were conducted according to OECD guidelines (2006) and standard methods (APHA-AWWA-WEF, 1998), the experimental data generated in the present study meets the quality criteria implied in the document. Therefore, Eq. 4.1, describing the relationship between hydrophobicity and marine algal toxicity has the potential to be used as a tool to carry out the preliminary risk assessments of polar narcotics to marine algae. These equations that employ Log *P* directly describe the mechanistic understanding of the toxicity like the proposed models in EU TGD (2003). It should be noted that these equations are valid in the Log *P* range 1.46 - 5.12.

4.1.2. QSTRs for marine alga *D. tertiolecta*

The Kolmogorov-Smirnov test revealed that marine algal toxicity of phenols to *D. tertiolecta* followed a normal distribution (p>0.05). The distribution of the marine algal toxicity data ($p\mathrm{T}_{48}$) is illustrated in Figure 4.2. Marine algal toxicity of 30 phenols to *D. tertiolecta* ($p\mathrm{T}_{48}$) used in QSTR modeling are presented in Table 4.3 together with the expected MOA, interaction type and selected descriptor values for each compound.

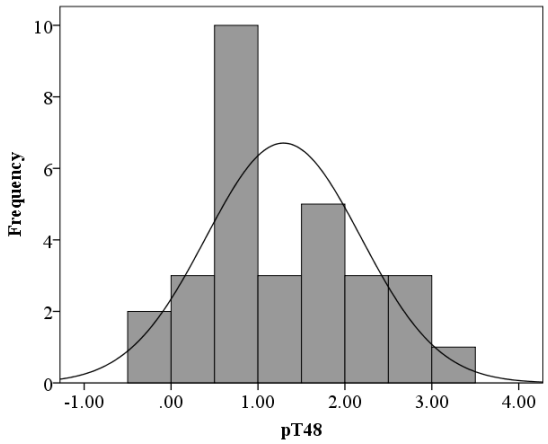


Figure 4.2. Histogram for the marine algal toxicity data ($p\mathrm{T}_{48}$)

Table 4.3. Mode of action (MOA), interaction type, 48 h marine algal toxicity and the descriptors used in the models

ID	Compound	MOA[a]	Interaction type[c]	pT_{48}	Mor24m	C_{ortho}	Mor24p	HATS7e	Mor18u	($k_1$max[a])
				Training set						
1	P	Polar narcosis	Non-covalent	-0.24(±0.03)[d]	0.068	-0.144	0.046	0	-0.50	9.08
2	2-CP	Polar narcosis	Non-covalent	0.44(±0.04)	0.031	-0.131	0.043	0	-0.43	9.14
4	4-CP	Polar narcosis	Non-covalent	0.81(±0.03)	0.107	-0.166	0.061	0	-0.47	9.05
5	2,3-DCP	Polar narcosis	Non-covalent	0.98(±0.04)	0.170	-0.135	0.070	0	-0.37	9.12
6	2,4-DCP	Polar narcosis	Non-covalent	1.23(±0.04)	0.103	-0.155	0.067	0	-0.40	9.10
7	2,5-DCP	Polar narcosis	Non-covalent	0.98(±0.01)	0.042	-0.133	0.066	0	-0.41	9.13
8	2,6-DCP	Polar narcosis	Non-covalent	0.43(±0.03)	0.004	-0.121	0.051	0	-0.44	9.17
10	3,5-DCP	Polar narcosis	Non-covalent	1.71(±0.05)	0.155	-0.151	0.108	0	-0.42	9.06
11	2,3,4-TCP	Polar narcosis	Non-covalent	1.88(±0.04)	0.329	-0.167	0.095	0	-0.31	9.08
12	2,3,5-TCP	Polar narcosis	Non-covalent	1.78(±0.04)	0.209	-0.139	0.103	0	-0.35	9.11
13	2,3,6-TCP	Polar narcosis	Non-covalent	0.80(±0.01)	0.144	-0.125	0.074	0	-0.39	9.15
15	2,4,6-TCP	Polar narcosis	Non-covalent	0.75(±0.02)	0.099	-0.147	0.081	0	-0.42	9.12
16	3,4,5-TCP	Polar narcosis[b]	Non-covalent	2.22(±0.02)	0.373	-0.189	0.123	0	-0.33	9.03
17	2,3,4,5-TeCP	Uncoupling	Non-covalent	2.59(±0.04)	0.448	-0.179	0.126	0	-0.27	9.08
18	2,3,4,6-TeCP	Uncoupling[b]	Non-covalent	2.29(±0.04)	0.318	-0.159	0.104	0	-0.35	9.11
19	2,3,5,6-TeCP	Uncoupling	Non-covalent	2.32(±0.03)	0.254	-0.130	0.102	0	-0.35	9.13
20	PCP	Uncoupling	Non-covalent	3.03(±0.04)	0.515	-0.171	0.135	0	-0.28	9.09
21	1,2,3-THB	Pro-electrophilic	Covalent	0.59(±0.02)	0.141	-0.197	0.071	0	-0.45	9.11
22	Hq	Pro-electrophilic	Covalent	2.56(±0.05)	0.097	0.059	0.049	0.042	-0.41	9.04
24	TetraClHq	Pro-redox cycling	Covalent	1.68(±0.03)	0.104	0.091	0.066	0.036	-0.43	9.09
26	4-ClCat	Pro-electrophilic	Covalent	0.86(±0.02)	0.123	-0.126	0.078	0	-0.47	9.08
27	3,5-DiClCat	Pro-electrophilic[b]	Covalent	1.19(±0.03)	0.182	-0.108	0.092	0	-0.49	9.09
28	Res	Polar narcosis	Non-covalent	-0.42(±0.02)	0.099	-0.239	0.056	0	-0.42	9.12
30	4,6-DiClRes	Polar narcosis	Non-covalent	0.51(±0.02)	0.146	-0.254	0.064	0	-0.43	9.13
				Test set						
3	3-CP	Polar narcosis	Non-covalent	0.91(±0.04)	0.124	-0.147	0.074	0	-0.46	9.07
9	3,4-DCP	Polar narcosis	Non-covalent	1.43(±0.04)	0.253	-0.178	0.088	0	-0.40	9.04
14	2,4,5-TCP	Polar narcosis	Non-covalent	1.99(±0.03)	0.199	-0.166	0.090	0	-0.35	9.09
23	ClHq	Pro-electrophilic	Covalent	2.59(±0.03)	0.151	0.071	0.065	0.039	-0.42	9.07
25	Cat	Pro-electrophilic	Covalent	0.73(±0.03)	0.107	-0.104	0.057	0	-0.47	9.09
29	4-ClRes	Polar narcosis	Non-covalent	0.16(±0.03)	0.133	-0.264	0.061	0	-0.45	9.15

[a] All MOAs were assigned according to Cronin et al. (2002). [b] Expected MOA - not included in the data set provided by Cronin et al. (2002). [c] Polar narcotics and respiratory uncouplers act through non-covalent mediated mechanisms while the reactive chemicals (e.g., pro-electrophiles, pro-redox cyclers, soft electrophiles, etc.) act through covalent interactions according to Schultz (1999). [d] Numbers in parenthesis refer to standard mean error of 48 h marine algal toxicity.

As discussed previously, it was observed that phenols tended to be slightly more toxic at 48 h exposure; therefore, the 48 h toxicity data was used as the dependent variable to be conservative in the modeling process. The toxicity values for this exposure duration varied uniformly over approximately 3.5 fold range (from -0.42 to 3.03 on a logarithmic scale).

The relationship between marine algal toxicity (pT_{48}) and hydrophobicity as described by Log *P* is illustrated in Figure 4.3.

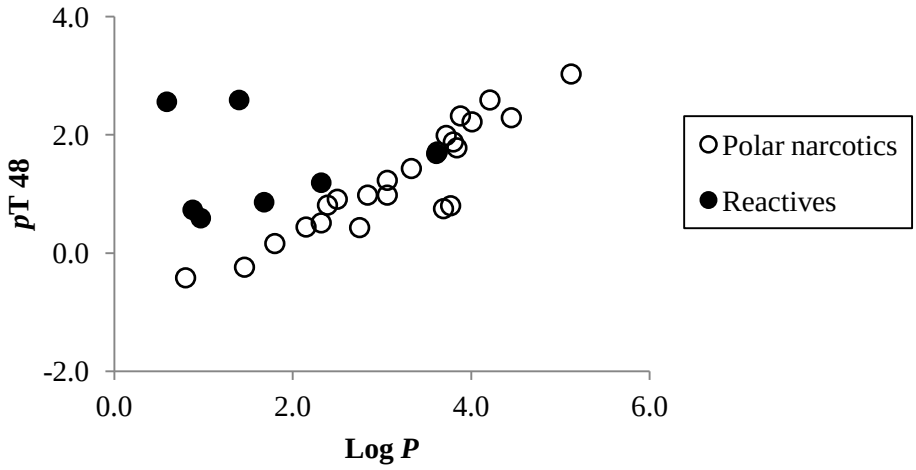


Figure 4.3. The relationship between marine algal toxicity of phenols and hydrophobicity as described by Log *P*.

The strong trend between hydrophobicity as described by Log *P* and toxicity for polar narcotics and respiratory uncouplers (collectively referred to as compounds acting through non-covalent mediated mechanisms), represented by open circles is obvious in Figure 4.3. Interestingly, the toxicity of tetrachlorohydroquinone (TetraClHq) had a reasonable agreement with Log *P*. As stated by Aptula et al. (2005b), this is probably because with increasing hydrophobicity the toxicity of electrophiles and pro-electrophiles converges on a narcosis mechanism in terms of both the observed physiological effects and the fit to the narcosis model.

The compounds expected to act by more reactive mechanisms (represented by black dots) were shown to have toxicity in excess of that predicted by hydrophobicity (Figure 4.3). This implicitly suggested that additional molecular descriptor(s) other than a

hydrophobicity term would be required to accurately predict the toxicity of phenols towards *D. tertiolecta*.

An initial modeling attempt was made using the pool of molecular descriptors including only 2D descriptors and Log *P*. However, all the resulting models had either low training set performance (i.e., $r^2 < 0.80$) or hydroquinones were found to be outliers. Following the removal of hydroquinones, the model presented in Eq. 4.2 resulted in a quality QSTR (which was developed using the training set in Table 4.3. without hydroquinones and its predictivity was assessed using compounds in the test set without hydroquinones):

$$pT_{48} = 0.63 \text{ Log } P + 1.725 \text{ MATS3p} - 0.38 \quad (4.2)$$

n_{tr}=22, r^2=0.89, r^2_{cv}= 0.80, $RMSE_{tr}$=0.30
n_{test}=5, Q^2_{F3}=0.82, R^2=0.80, $RMSE_{te}$=0.37

In Eq. 4.2, the Log *P* describes hydrophobicity while MATS3p is a 2D autocorrelation descriptor weighted by atomic polarizability (originally named as Moran autocorrelation lag three weighted by atomic polarizability) obtained from the Dragon 5.4 software. The positive contribution of this descriptor suggests that algal toxicity increases with increasing atomic polarizability of phenols. Although this two-parameter QSTR produced highly satisfactory results, it was not capable of explaining the toxicity of hydroquinones. Therefore, to model the toxicity of phenols including hydroquinones to *D. tertiolecta*, 3D descriptors were included in the modeling process in search for global models which preferably did not produce any compound as outlier.

4.1.3. Global models

The European Centre for Ecotoxicology and Toxicology of Chemicals has published several reports (ECETOC, 1993; ECETOC, 2001) to highlight the need for marine toxicity data and recently in the workshop on the probabilistic approaches to be used in marine risk assessment (ECETOC, 2008), experts have recognized the potential use of QSTRs in efforts to fill the data gap present in this domain of ecotoxicology. One of the

recommendations of the workshop was to develop models for compounds having similar mode of actions (i.e., local models). For regulatory usage, it is clearly important to develop and understand local models because this enables confidence to be better assigned to the predictions when these models are used in a real life scenario (Enoch et al., 2008). However, it should be kept in mind that the knowledge on the MOA of chemicals varies considerably (Spycher et al., 2008) and assignment of the correct MOA is a difficult task (Schultz, 1999). Certain mechanisms may even be specific for the given species and these can be related to the metabolism and distribution of the compounds *in vivo* (Aptula et al., 2005b). Therefore, as stated by Schultz (1999), the models that accurately predict the toxicity of chemicals without first identifying their toxic mechanisms (i.e., global models), are highly desirable.

Efforts to model the data set without classifying the chemicals into MOAs resulted in three global models which displayed acceptable training ($r^2 \geq 0.80$, $r^2_{cv} \geq 0.70$) and test set ($Q^2_{F3} \geq 0.70$ and $R^2 > 0.60$) statistics and these are listed in Table 4.4 together with the statistics of the consensus MLR model. The models in Table 4.4 were also subjected to the predictivity criteria proposed by Golbraikh et al. (2003) as summarized in equations (3.4-3.7) and also to *Y*-randomization (Table 4.5).

It was found that r_r^2 of the randomized models were lower than the corresponding r^2 of the non-randomized (i.e., original) model and the R_p^2 values were higher than 0.5. This finding indicated that the obtained relationships were not due to chance. At the end of the rigorous validation process, we were convinced that the proposed models were robust, predictive and not chance correlations.

It should be noted that the consensus model integrating all validated individual models was found to be the most externally predictive. This finding indicates that consensus modeling approach can be used to predict the toxicity of phenols towards *D. tertiolecta*. Previously, Zhu et al. (2008) reached a similar conclusion in their efforts to offer global QSTRs to predict the toxicity of a large set of phenols towards *T. pyriformis*. They noted that the consensus models offer better predictivity than the individual models, which was similar to our finding for the relatively small marine algal toxicity data set.

Table 4.4. Marine algal QSTRs that displayed acceptable training and test set statistics

Model ID	MLR Models	Training set				Test set			
		n	r^2	r^2_{cv}	$RMSE_{tr}$	*n*	R^2	Q^2_{F3}	$RMSE_{te}$
1	pT_{48h}= 6.18 Mor24m + 6.27 C_2 +1.06	24	0.82	0.74	0.38	6	0.80	0.85	0.34
2	pT_{48h} = 31.10 Mor24p + 44.41 HATS7e -1.36	24	0.83	0.71	0.36	6	0.86	0.88	0.31
3	pT_{48h} = 11.21 Mor18u + 4.37 C_2 – 8.66 k_1max[a] + 85.13	24	0.80	0.73	0.39	6	0.90	0.92	0.26
MLR Consensus Model[*]		24	0.88	0.86	0.31	6	0.95	0.94	0.23

[*] Average of the three models (Model 1, 2, and 3)

Table 4.5. Additional external validation criteria and *Y*-randomization for marine algal QSTRs

Model ID	Criteria by Golbraikh et al. (2003)					*Y*-randomization	
	$\frac{R^2 - R_0^2}{R^2} < 0.1$	$0.85 \leq K \leq 1.15$	$\frac{R^2 - R_0'^2}{R^2} < 0.1$	$0.85 \leq K' \leq 1.15$	$\lvert R_0^2 - R_0'^2 \rvert \leq 0.3$	r_r^2	$R_p^2 > 0.5$
1	0.046	0.893	0.001	1.068	0.038	0.09	0.70
2	0.023	0.872	0.001	1.109	0.020	0.09	0.71
3	0.007	0.913	0.025	1.067	0.016	0.12	0.66
MLR Consensus	0.005	0.892	0.001	1.105	0.004	-	-

4.1.4. Mechanistic interpretation of the global models

The descriptors that appeared in the linear QSTRs (Table 4.4) involve Mor24m (3D-MoRSE – signal 24 / weighted by atomic masses), Mor24p (3D-MoRSE – signal 24 / weighted by atomic polarizabilities) and Mor18u (3D-MoRSE – signal 18 / unweighted) obtained from the Dragon 5.4. Another Dragon descriptor, namely the HATS7e (Leverage weighted autocorrelation of lag 7 / weighted by atomic Sanderson electronegativities) is a GEometry, Topology, and Atom-Weights AssemblY (GETAWAY) descriptor which encode both the geometrical information provided by the influence molecular matrix and the topological information provided by the molecular graph, weighted by chemical information encoded in selected atomic weightings (Consonni et al., 2002). The descriptors that were obtained from the CODESSA software were C_2 (originally named as MOPAC partial charges for atom # 0000002 - C) and k_1max^a, which indicate the partial charge on the second atom (carbon) located at the (2-) position relative to the functional group (i.e., the OH moiety) and the maximum atomic force constant, respectively. For the latter, Katritzky et al. (1995b) stated that this descriptor can be associated with intermolecular interactions.

The commonality of the three QSTRs developed in this study is the presence of 3D-MoRSE descriptors (Table 4.4). These descriptors provide information of a compound from its three-dimensional structure using a molecular transform derived from an equation used in electron diffraction studies and by taking several atomic properties into account they have the potential to encode highly flexible representations of a molecule (Mercader and Pomilio, 2010). The presence of Mor24m (3D-MoRSE signal 24/ weighted by atomic masses) and Mor24p (3D-MoRSE signal 24/ weighted by atomic polarizabilities) in two distinct QSTRs (model 1 and model 2 in Table 4.4) suggests that biological activity has a significant dependence on the size and polarizability of the molecules. On the other hand, the presence of Mor18u (3D-MoRSE – signal 18 / unweighted), which is neither influenced by the atomic polarizability nor by the atomic mass of phenols, suggests that the 3D structure of the molecules might have an impact on their toxicity as well.

In an effort to better understand the role of 3D-MoRSE descriptors and interpret the developed MLR models mechanistically, the whole dataset was decomposed into several

subsets according to a) the structure of the compounds (i.e., simple phenols and polyphenols) and b) the nature of mechanisms (i.e., non-covalent and covalent) as described by Schultz (1999). The subset analyses using 3D-MoRSE descriptors were also carried out for subgroups excluding hydroquinones (Table 4.6).

Table 4.6. Correlation of 3D-MoRSE descriptors to the toxicity of phenols to *D. tertiolecta*

Subset ID	Subset specification	*n*	Mor24m		Mor24p		Mor18u	
			r^2	r^2_{cv}	r^2	r^2_{cv}	r^2	r^2_{cv}
1	Phenol + Chlorophenols	20	**0.80**[a]	**0.76**	**0.88**	**0.85**	**0.78**	**0.73**
2	Polyphenols	10	0.00	n.t.	0.00	n.t.	0.06	n.t.
3	Polyphenols (except hydroquinones)	7	0.43	0.03	0.59	0.30	0.74	0.46
4	Non-covalent	23	0.72	0.68	**0.85**	**0.83**	0.73	0.68
5	Covalent	7	0.02	n.t.	0.20	n.t.	0.56	0.22
6	Covalent (except hydroquinones)	4	0.51	0.00	0.64	0.05	**0.97**	**0.92**
	All phenols	30	0.50	0.46	0.49	0.42	0.49	0.43
	All phenols (except hydroquinones)	27	0.72	0.69	**0.84**	**0.82**	0.65	0.59

n.t. not tested due to low r^2.
[a] relationships with r^2 and $r^2_{cv} \geq 0.75$ are highlighted in bold.

The analyses revealed that the 3D-MoRSE descriptors were able to predict the toxicity of chlorophenols (Table 4.6, Subset ID 1) whereas they failed to do so for polyphenols (Table 4.6, Subset ID 2) even when hydroquinones were removed (Table 4.6, Subset ID 3). This is not unexpected because polyphenols are involved in different reactivity mechanisms and a structure based-classification would be futile in efforts to predict the toxicity of poly-hydroxylated aromatics. However, since chlorophenols have structural as well as mechanistic similarity, all the 3D-MoRSE descriptors, particularly Mor24p, was found to be in excellent agreement with the toxicity of this class of compounds (Table 4.6, Subset ID 1: r^2=0.88, r^2_{cv}=0.85), as well as with compounds eliciting toxic effects through non-covalent mediated interactions (Table 4.6, Subset ID 4: r^2=0.85, r^2_{cv}=0.83). Closer inspection of Mor24p (Table 4.3) revealed that chlorophenols containing chlorine atoms at 2- and 6- positions have significantly lower values compared to the congeners with the same number of chlorine atoms. Note that the *ortho* substituted chlorophenols were found to be less toxic than the *meta* and *para* substituted congeners also (Table 4.3). As stated by Boyd et al. (2001), the lower toxicity of *ortho* substituted chlorophenols has been ascribed to the shielding of the OH group by chlorine atom(s). Mor24m was also found to be useful in predicting the toxicity of chlorophenols (Table 4.6, Subset ID 1: r^2=0.80, r^2_{cv}=0.76); however, considering the superiority of Mor24p over

Mor24m, it might be inferred that 3D information weighted by atomic polarizability might be more important in the toxicity of chlorophenols to *D. tertiolecta* as compared to that weighted by atomic mass. It was also interesting to note that Mor24p was alone able to explain 84% of the variance in the data set when hydroquinones were removed from analysis (Table 4.6, Figure 4.4). The high correlation of Mor24p with phenols except hydroquinones implicitly suggests that the presence of the additional descriptor (HATS7e) in the developed QSTR (Model 2, Table 4.4) is mainly to account for the toxicity of hydroquinones. In line with this expectation, closer inspection of HATS7e revealed that this descriptor has values other than zero only for hydroquinones (Table 4.3). The concomitance of Mor24m and C_2 can also be interpreted similarly for model 1 (Table 4.4). Inspection of C_2 revealed distinct values for the three hydroquinones in that only these compounds had positive partial charges in the second carbon atom (Table 4.3). Constitution of model 3 (Table 4.4), on the other hand, differs from the other models in that there are two additional descriptors (i.e., C_2 and k_1max^a). Note that Mor18u was shown to be correlated very strongly with the toxicity of phenols acting through covalent interactions except hydroquinones (Table 4.6, Subset ID 6: r^2=0.97, r^2_{cv}=0.92). Considering the relatively good agreement of this descriptor with compounds acting through non-covalent interactions also (Table 4.6, Subset ID 4: r^2=0.73, r^2_{cv}=0.68), the presence of C_2 and k_1max^a can be said to explain the remaining variance in the toxicity, particularly that inflicted by compounds acting through covalent mediated mechanisms (Table 4.3).

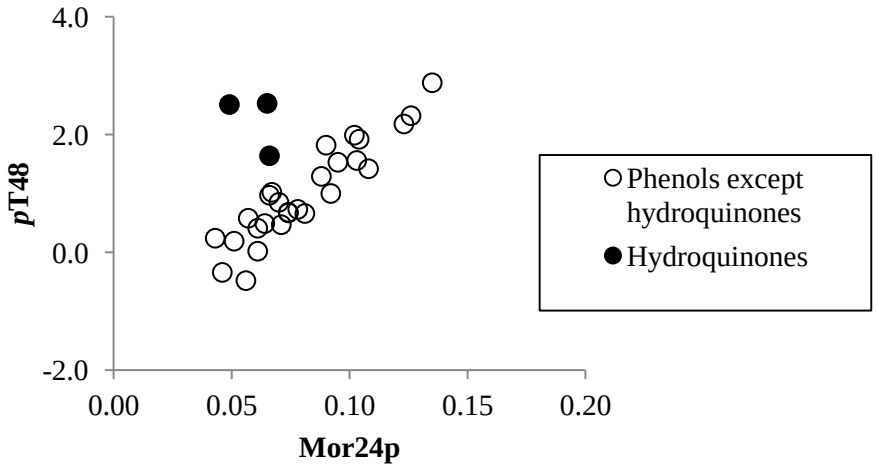


Figure 4.4. The relationship between 48 h marine algal toxicity and Mor24p

From an organic chemistry perspective, hydroquinones can undergo one-electron oxidation to semi-quinone, and being a radical anion the semi-quinone can be further oxidized to benzoquinone. Alternatively, hydroquinone can be oxidized directly to benzoquinone in a two-electron transfer reaction (Aptula et al., 2005b). As stated by Aptula et al. (2009), hydroquinone is usually considered to be a pro-electrophilic sensitizer that can be oxidized either abiotically or enzymatically to benzoquinone which is a chemically reactive Michael type acceptor electrophile. The researchers also noted that benzoquinone itself is recognized as an extreme sensitizer, and guinea pig cross-challenges pointed out that hydroquinone sensitizes via benzoquinone. Satoh et al. (2009) also indicated that hydroquinones are not themselves electrophilic and require oxidative conversion to their quinone forms to be involved in electrophilic processes. On the basis of the explanations above, the reason behind the excess toxicity of hydroquinones (particularly hydroquinone and chlorohydroquinone) towards *D. tertiolecta* possibly lies in the transformation of these compounds into their more electrophilic quinone forms. As for the lower toxicity of tetra-substituted hydroquinone, TetraClHq, a parallel observation regarding the reduced toxicity to *T. pyriformis* with increasing substitution (other than mono-substitution) on the aromatic ring was also observed by Aptula et al. (2005b), particularly for methyl and methoxy substituted hydroquinones. They interpreted this as a result of different reaction possibilities that these additional groups might introduce. In another study where they investigated the toxicity of quinones to *T. pyriformis*, Schultz et al. (1997) reported that the toxicity of benzoquinone and chlorobenzoquinone to *T. pyrifomis* was close to each other while the toxicity of tetrachlorobenzoquinone was significantly lower. The researchers stated that ring substitution of 1,4-benzoquinone resulted in a graded toxic response and mono-chloro substitution (i.e., chlorobenzoquinone) had little effect on the toxic potency compared to benzoquinone while tetra-chloro substitution (i.e., tetrachlorobenzoquinone) sharply reduced the toxicity. Note that the toxicity of hydroquinone (pT_{48}=2.56) and chlorohydroquinone (pT_{48}=2.59) to *D. tertiolecta* was also close to each other and significantly higher than that of tetrachlorohydroquinone (pT_{48}=1.68). The chloro-substitution versus toxicity pattern observed for the toxic potency of benzoquinone(s) towards *T. pyriformis* is similar to what we observed in this study for the toxicity of hydroquinones towards *D. tertiolecta*. Although the difference in the biological constitution of protozoa and algae makes it difficult to indicate a precise opinion, the toxicity versus chloro substitution pattern

mentioned above implicitly suggests that the hydroquinones displayed toxic characteristics of their benzoquinone counterparts towards *D. tertiolecta*.

From a modeling perspective, the information regarding the excess toxicity of hydroquinones was best captured by HATS7e and C_2 in the QSTRs. However, the fact that both descriptors were able to discriminate excess hydroquinone toxicity should not be interpreted as the ability of these descriptors in describing the reactivity of phenols. The information captured by both descriptors, particularly by C_2, can be interpreted as the propensity of a pro-electrophile to undergo oxidation reactions and transform into its more electrophilic form or, alternatively, the propensity of a compound to interact with nucleophilic target sites at this particular carbon atom. Additionally, the low C_2 values of resorcinols (Table 4.3) imply that this class of compounds may not be active in electrophilic/nucleophilic processes in the marine algal test system. This interpretation is also in line with the organic chemistry principle that resorcinols are not involved in oxidation reactions unless activated by additional groups (Aptula et al., 2005b; Lyakhovich et al., 2006).

The difficulties in modeling the toxicity of compounds that may act either directly or indirectly as electrophiles have been reported in various studies (Cronin et al., 2002; Enoch et al., 2008). The difficulty in modeling the toxicity of such compounds might be due to the fact that aromatic compounds are hydroxylated by intracellular monooxygenases *in vivo*, and it is, therefore, difficult to discriminate the effect of initial compound and its metabolites (Lyakhovich et al., 2006). Additionally, from a statistical perspective, the dominance in the data set of compounds acting through similar MOAs (i.e., compounds acting through non-covalent MOAs like polar narcosis and respiratory uncoupling) results in the selection of a descriptor, or a set of descriptors, that best explain(s) these mechanisms (Enoch et al., 2008). However, compounds acting through other MOAs and consequently the associated descriptors may not be equally represented in the models purely due to issues related to the data set bias. For instance, the marine algal data set presented in this study includes chlorophenols and resorcinols, which act through non-covalent mediated interactions; therefore, the marine algal QSTRs can be expected to explain the toxicity of these compounds first to be able to explain most of the variance. Indeed, the developed QSTRs included 3D-MoRSE descriptors which were able to explain

a great part of the variance in the toxicity of non-covalent mediated toxicity (Table 4.6, Subset ID 4). The remaining descriptors were likely to be selected in the models mainly to account for the variance in the toxicity of compounds acting through covalent interactions, particularly hydroquinones.

4.1.5. Applicability domain of the global models

The applicability domain of each model was constructed using the plot of standardized residuals against leverages (Figure 4.5).

As can be seen from Figure 4.5, all compounds had standardized residuals < 2.5 indicating that there was not a response outlier (Y outliers) in the developed QSARs. The cut-off hat value for model 1, model 2, and model 3 were found to be 0.375, 0.375, and 0.5, respectively. On the basis of these cut-off values, tetrachlorohydroquinone (ID=24) was detected as an X outlier for model 1 (Figure 4.5a). For model 2, both tetrachlorohydroquinone and hydroquinone (ID = 22) were X outliers (Figure 4.5b); while there was not any X outlier for model 3 (Figure 4.5c).

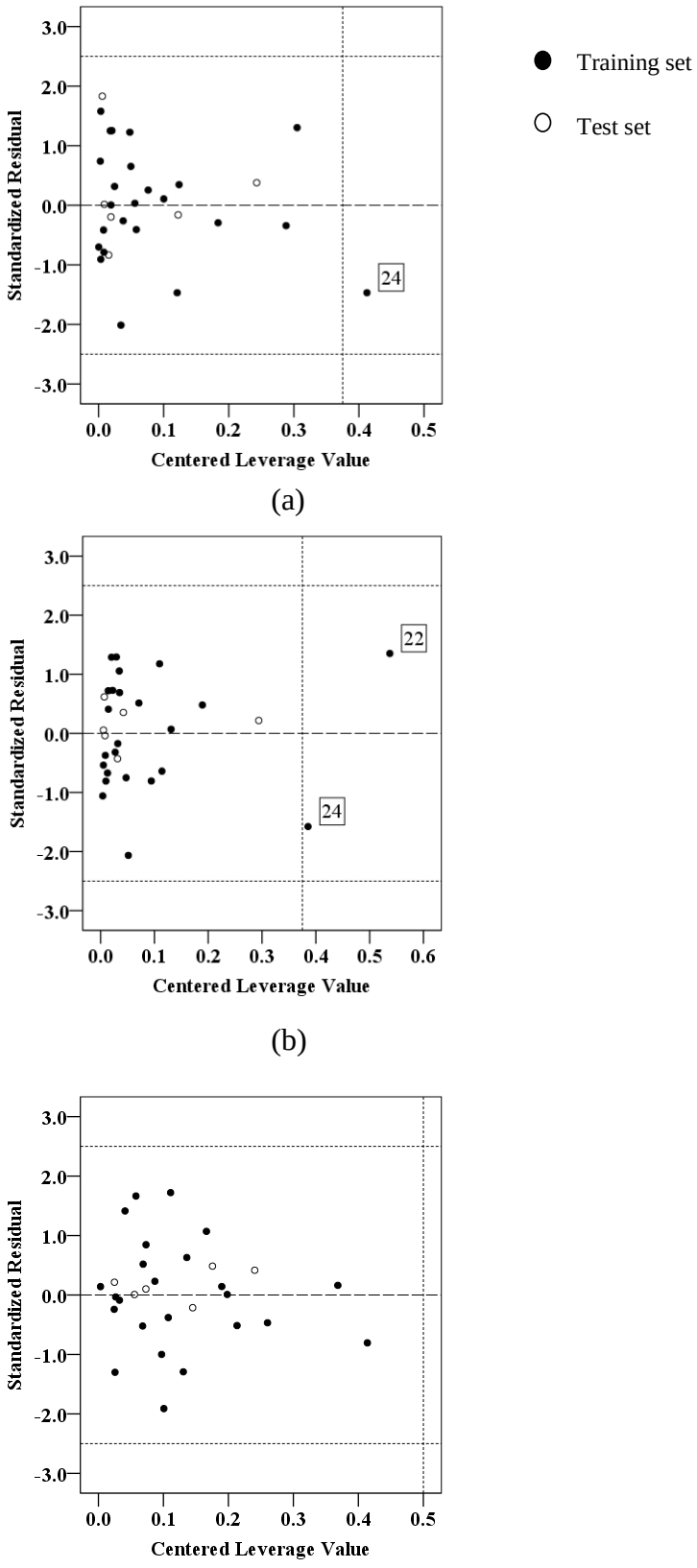


Figure 4.5. Applicability domains of a) model 1 b) model 2 and c) model 3 developed for *Dunaliella tertiolecta*

4.1.6. Counter Propagation Artificial Neural Networks (CP ANNs)

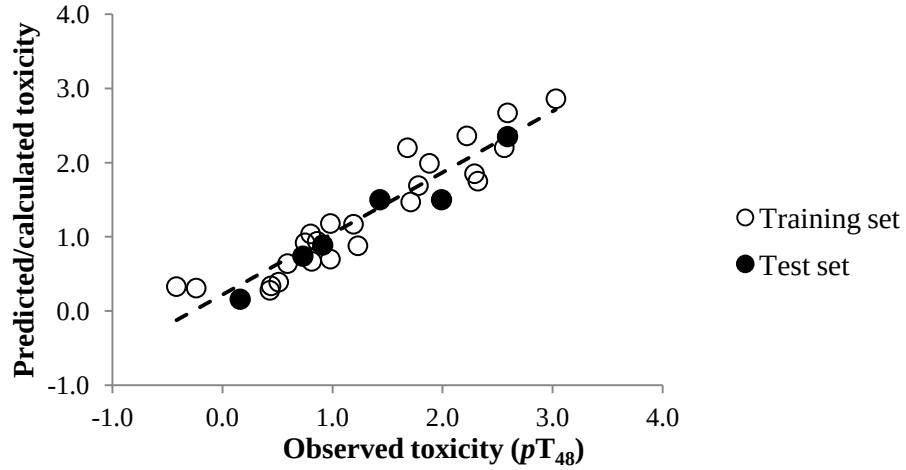
It can be assumed that the consensus MLR model (Table 4.4) carries significant toxicological information that is encoded in the descriptors that appeared in the three MLR models. In an effort to understand how these significant descriptors would contribute to a non-linear model, an initial attempt was made to construct a CP ANN model using all six molecular descriptors that appeared in the proposed linear models (Mor24m, C_2, Mor24p, HATS7e, Mor18u, and k_1max^a). However, it was found that Mor24m and Mor24p were inter-correlated; therefore, the former was removed from analysis while the latter was kept in as it had higher correlation to toxicity. Likewise, C_2 and HATS7e were also found to be inter-correlated and the latter was removed from modeling due to the presence of zero values for most of the compounds.

With the introduction of four descriptors as input variables, the four-parameter CP ANN modeling was performed employing three different network architectures (5x5, 6x6 and 7x7) as well as a range of learning iterations (10-400 iterations) to cover all investigated objects. The modeling was performed on the training set (24 compounds), employing the leave-one-out cross-validation procedure for optimization of the model parameters. The best CP ANN model was selected according to the highest r^2_{cv} value which was obtained from a network of 6x6 dimensions and 350 epochs.

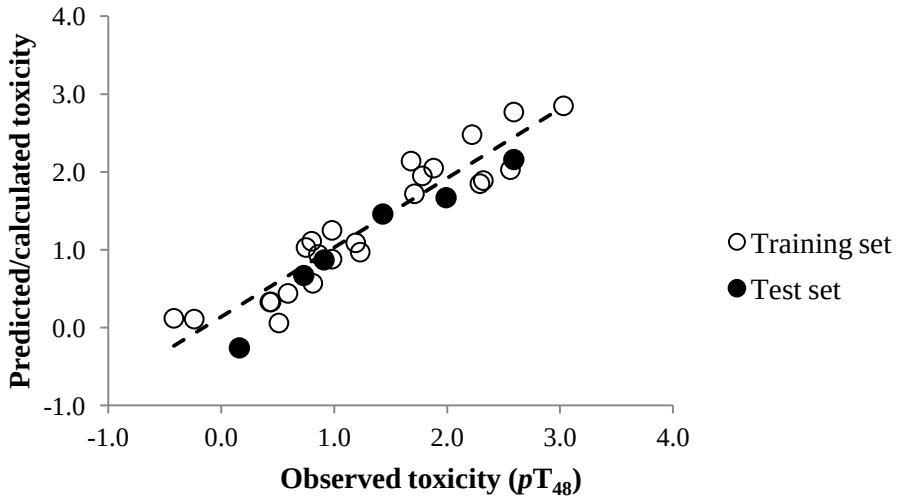
The results revealed that the statistical parameters of the four-parameter CP ANN model were found to be comparable to the consensus MLR model (Table 4.7, Figure 4.6).

Table 4.7. Statistical comparison of the MLR consensus and CP ANN models

Model	Training Set				Test set		
	n	r^2	r^2_{cv}	$RMSE_{tr}$	*n*	Q^2_{F3}	$RMSE_{te}$
MLR Consensus	24	0.88	0.86	0.31	6	0.94	0.23
CP ANN	24	0.92	0.73	0.26	6	0.93	0.25



(a)



(b)

Figure 4.6. The observed versus predicted/calculated toxicity for a) MLR and b) CP ANN models

These results apparently show the good predictive ability of the CP ANN model for the prediction of the marine algal toxicity of phenols to *D. tertiolecta*. Clearly, the four-parameter CP ANN model should not be read strictly as a non-linear counterpart of the consensus MLR model. Yet, it provides useful information regarding the descriptors or descriptor combinations used in the present study. Additionally, the CP ANN model

enabled a successful clustering of the phenols with similar biological toxicity as well as interaction type over the entire network (Figure 4.7e).

The visual analysis of the output layers for the obtained CP ANN model (Figure 4.7) highlighted the good agreement between the weight maps for the 3D-MoRSE descriptors and toxicity. Mor24p (Figure 4.7c) and Mor18u (Figure 4.7d) displayed very similar color distribution to that of the biological toxicity (Figure 4.7e) which accounts for a good positive correlation for almost all investigated compounds, and especially for the chlorophenols. On the other hand, the weights for C_2 (Figure 4.7a) and k_1max^a (Figure 4.7b), respectively, follow a different color distribution compared to the biological toxicity map (Figure 4.7e).

C_2 molecular descriptor discriminates two classes: hydroquinones (ID = 22, 23, 24) and resorcinols (ID = 28, 29, 30); while k_1max^a molecular descriptor clearly separates the catechols (ID = 25, 26, 27). Note that the descriptor's values in Figure 4.7 (a, b, c, and d) are normalized between 0.00 (minimum value) and 1.00 (maximum value). The output layer for toxicity (pT_{48}) is presented as E, where the values 0.05 and 2.81 correspond to minimum and maximum toxicities predicted by the CP ANN model, respectively. The numbers in the plot (e) correspond to the positions of the compounds from the training (white) and test set (black) in the *top-map* of the Kohonen layer.

4.1.7. External predictivity of the global models

The proposed QSTRs were further tested in external predictivity using a heterogeneous data set obtained from Aruoja et al. (2011). The external prediction performance of the proposed models is provided in Table 4.8. The descriptor values for the external set compounds are provided in Appendix D (Table D1).

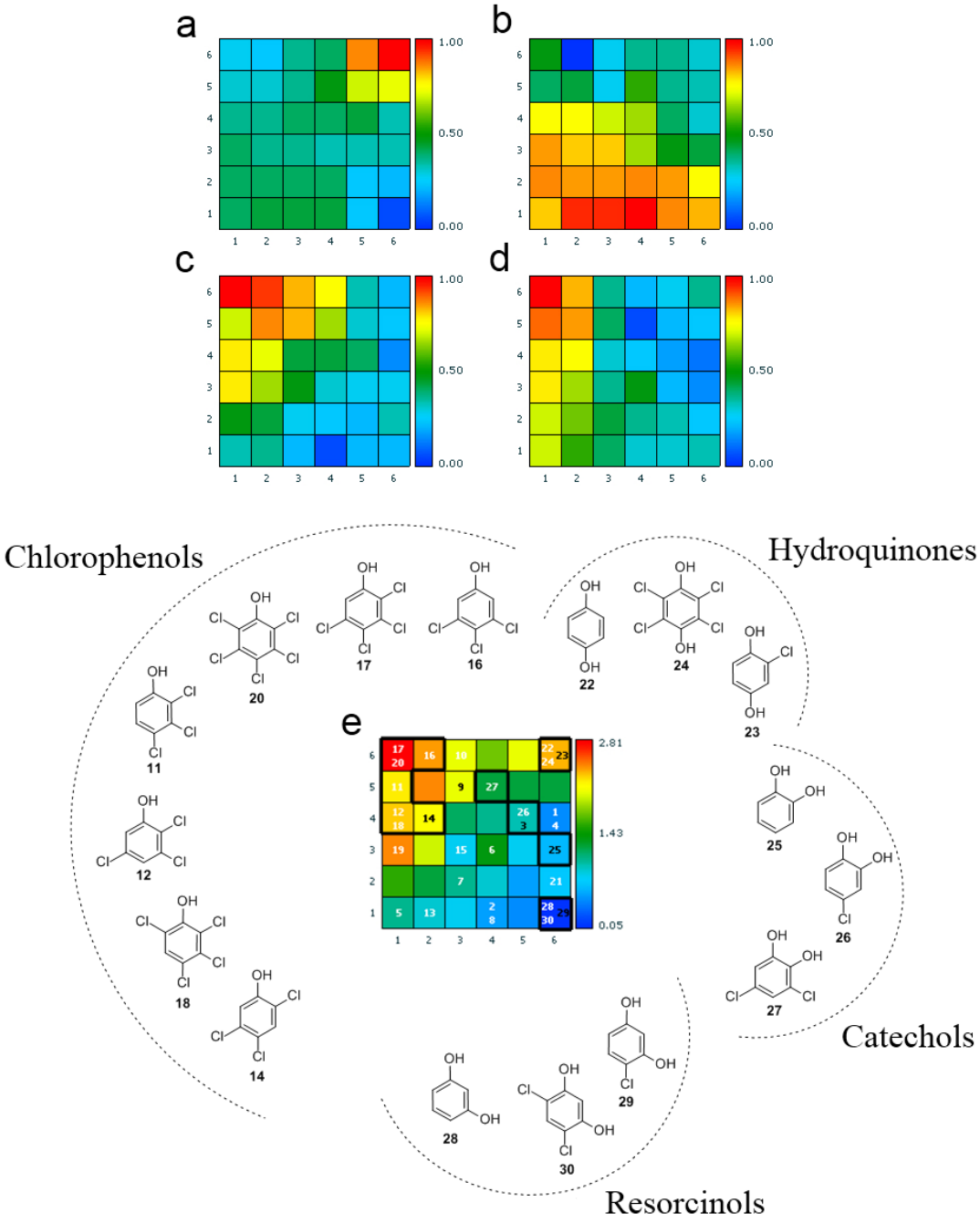


Figure 4.7. Output layers of the selected four-parameter CP ANN model. **a)** C_2; **b)** k_1max^a; **c)** Mor24p; **d)** Mor18u; **e)** pT_{48h} and descriptor-based clustering of phenols over the CP ANN network (Compound ID in white and black refer to training and test set compounds, respectively).

Table 4.8. External prediction performance of the marine algal QSTRs

Model	Number of external set compounds	Compounds out of AD (h>h*)	Number of response outliers	n[a]	Q^2_{F3}	$RMSE_{ext}$	% coverage[b]
Model 1	58	0	1	57	0.73	0.47	100
Model 2	58	20	3	35	0.69	0.49	65
Model 3	58	38	4	16	0.69	0.50	34

a: Number of compounds put into analysis after the removal of response outliers and compounds out of AD which were poorly predicted by the model.
b: Ratio of the number of external set compounds within the model AD to the total number of compounds in the external set, expressed as percentage.

Table 4.9. Additional external prediction criteria proposed by Golbraikh et al. (2003)

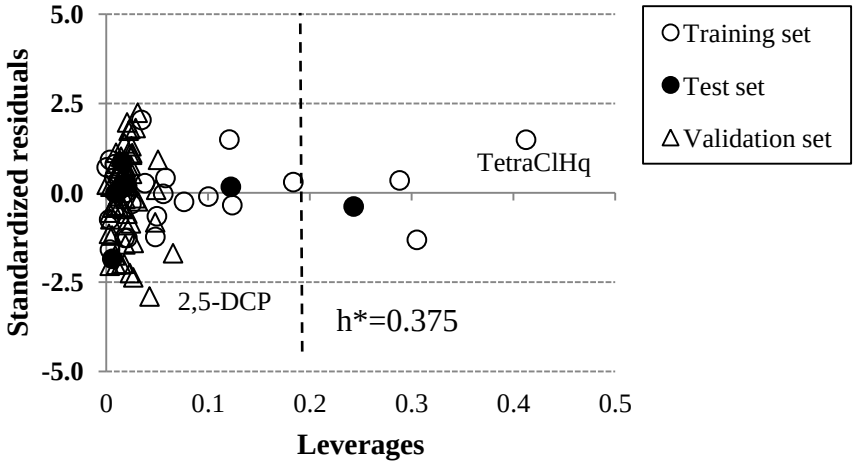
Model	**Criteria by Golbraikh et al. (2003)**				
	$\frac{R^2 - R_0^2}{R^2} < 0.1$	$0.85 \leq K \leq 1.15$	$\frac{R^2 - R_0'^2}{R^2} < 0.1$	$0.85 \leq K' \leq 1.15$	$\left\|R_0^2 - R_0'^2\right\| \leq 0.3$
Model 1	0.06	1.04	0.05	0.90	0.01
Model 2	0.07	**1.20**[a]	0.01	**0.78**	0.05
Model 3	0.04	1.05	**0.12**	0.88	0.07

a: Criteria that were not met by the model were highlighted in bold.

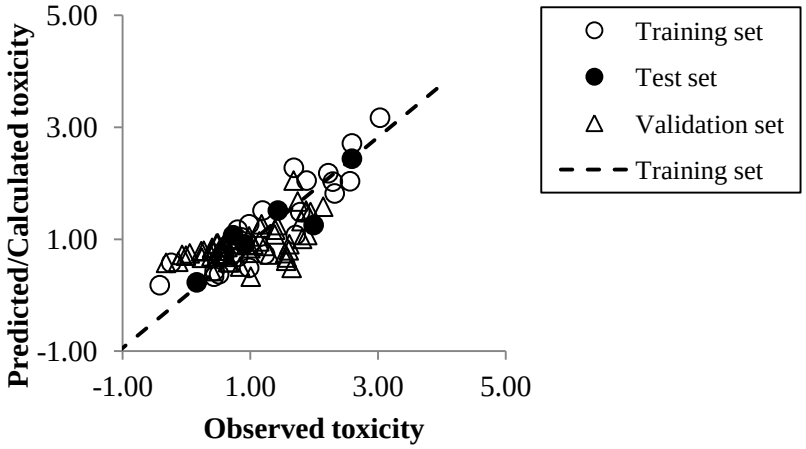
As indicated by Kar and Roy (2010), Q^2 values higher than 0.5 are considered acceptable in terms of a model's external predictivity. Accordingly, following the removal of one response outlier, namely, 2,5-DCP, the predictive performance of model 1 (Table 4.8) was found to be very good (Q^2_{F3}=0.73), especially considering its perfect coverage of the external data set. On the other hand, model 2 and in particular model 3 had narrow applicability domains which were reflected in low coverage of the external dataset by these models (%65 and %34, respectively, Table 4.8). Further evaluation of the proposed QSTRs according to the external prediction criteria put forward by Golbraikh et al. (2003) revealed that while model 1 met all the criteria, model 2 and model 3 failed to concur with some of the requirements (Table 4.9). The 4-parameter CP ANN, which was built using all the descriptors appearing in model 3 plus another one from model 2 (i.e., Mor24p), had also low predictivity and coverage (results not shown).

Of the proposed models, model 1 was highlighted because of its overall performance. The Williams plot and the observed versus calculated/predicted toxicity for model 1 are illustrated in Figure 4.8a and Figure 4.8b, respectively. The predicted toxicities and leverages by model 1 as well as for model 2 and model 3 for the external set compounds are provided in Appendix D (Table D2).

As mentioned earlier and can be seen from Figure 4.8a, tetrachlorohydroquinone was found to be a structurally influential compound for model 1, although the visual inspection revealed that it was a good leverage compound. Additionally, 2,5-DCP in the external set was found to be more toxic than predicted by model 1 towards *P. subcapitata*. Following the exclusion of 2,5-DCP from analysis, the external predictivity of model 1 was found to be highly satisfactory (Table 4.8, Figure 4.8b). It should also be noted that the applicability domain of model 1 was wide enough to cover phenols with different substituents than chlorine (i.e., methyl and ethyl) as well as another class of organic compounds (i.e., anilines).



(a)



(b)

Figure 4.8. For model 1, (a) the Williams plot (h* indicates the warning hat value), and (b) the plot of observed toxicity versus calculated/predicted toxicity for training, test and validation sets

4.2. Toxicity of selected phenols to freshwater alga *Chlorella vulgaris*

All bioassays conducted with freshwater alga *C. vulgaris* concurred with the acceptability criteria recommended by the OECD (2006). At the end of 96 h, the algal biomass increased by approximately 20-fold (corresponding to a mean control growth rate of 0.65 (±0.08) d^{-1} at the end of 96 h). The coefficient of variation within the controls was found to be 5%. The *p*H in the beginning and at the end of the bioassays was 6.36 (±0.02) and 6.56 (±0.05), respectively.

The GC analysis revealed that the tested concentration (100 mg/L) of 2-CP decreased by 5% at the end of 96 h. It should be noted that since none of the test concentrations changed by more than 20%, the nominal concentrations were used to calculate the toxic potency of each compound as recommended by the OECD (2006).

The experimental IC_{50} and IC_{20} values and the associated confidence intervals obtained at the end of 96 h based upon specific growth rate (SGR) and yield are presented in Table 4.10 together with the toxic class, NOEC and LOEC of each test chemical.

Table 4.10. 50% and 20% inhibitory concentrations (IC_{50} and IC_{20}) calculated at the end of 48, 72 and 96 h based on three different methods using yield and specific growth rate (SGR) calculations, no-observed-effect concentration (NOEC), lowest-observed-effect concentration (LOEC), toxic class of phenols for *C. vulgaris*

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class[b]	NOEC-LOEC[c] (mg/L)
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
P	Yield	ICp	313.3 (293.9-336.9)[a]	175.4 (48.2-317.0)	298.1 (210.4-352.9)	107.5 (8.2-409.6)	298.4 (262.0-333.5)	99.5 (45.7-423.6)	Not classified	<120/120
		Weibull	309.7 (211.2-376.1)	196.1 (57.5-257.8)	281.3 (181.9-346.7)	170.7 (44.4-231.1)	285.9 (208.3-338.5)	191.1 (74.4-242.6)		
		Polynomial Regression	296.4 (231.8-336.1)	127.9 (102.8-154.7)	294.3 (199.6-367.1)	112.2 (71.3-193.1)	272.1 (214.3-306.5)	108.9 (81.5-145.5)		
	SGR	ICp	389.0 (355.9-415.3)	270.8 (253.9-292.8)	383.4 (334.7-418.5)	270.2 (149.6-319.1)	374.9 (362.7-386.6)	276.3 (246.1-304.7)	Not classified	120/240
		Weibull	392.4 (300.2-473.6)	267.4 (92.8-331.5)	379.3 (302.6-435.9)	277.2 (130.7-331.7)	369.6 (327.3-408.2)	300.0 (200.9-335.3)		
		Polynomial Regression	412.1 (391.9-443.8)	207.6 (183.0-240.4)	395.0 (334.9-473.8)	193.6 (146.9-271.3)	372.3 (371.7-376.1)	267.1 (251.5-283.2)		
2-CP	Yield	ICp	65.5 (53.9-76.2)	27.0 (13.7-41.6)	59.7 (25.5-86.0)	15.7 (1.8-60.7)	62.4 (34.2-79.3)	13.2 (5.0-37.4)	Harmful	<15/15
		Weibull	61.6 (38.7-89.1)	27.5 (6.6-42.3)	50.7 (28.5-77.4)	20.2 (3.1-33.4)	48.6 (26.9-74.6)	19.1 (2.7-31.9)		
		Polynomial Regression	62.6 (53.4-70.8)	24.0 (18.8-30.0)	50.9 (38.4-73.2)	18.3 (7.7-37.0)	48.3 (41.7-60.3)	17.1 (10.4-26.6)		
	SGR	ICp	95.3 (88.2-104.7)	48.2 (39.8-56.3)	90.2 (79.8-101.8)	48.1 (9.7-75.4)	86.3 (79.6-93.0)	55.4 (26.6-70.1)	Harmful	15/30
		Weibull	92.3 (64.6-125.1)	48.7 (15.7-68.2)	86.0 (61.4-111.8)	48.3 (17.5-65.8)	82.4 (65.3-98.2)	54.2 (32.1-67.6)		
		Polynomial Regression	96.6 (86.8-108.1)	39.4 (33.9-46.3)	86.0 (72.0-105.7)	35.6 (26.4-49.0)	77.3 (67.7-89.7)	33.3 (27.9-40.4)		
3-CP	Yield	ICp	23.7 (21.6-25.3)	14.7 (8.6-18.2)	26.7 (23.5-31.0)	9.0 (2.2-14.7)	34.7 (30.1-39.0)	4.9 (2.9-12.6)	Harmful	<5/5
		Weibull	24.2 (15.6-33.1)	12.3 (2.7-17.9)	25.3 (13.2-37.5)	10.0 (1.4-17.0)	26.0 (11.7-40.6)	9.1 (0.8-16.8)		
		Polynomial Regression	25.5 (23.0-27.8)	10.4 (8.2-12.6)	26.6 (23.4-30.2)	9.2 (6.8-12.0)	28.9 (25.0-33.5)	8.5 (5.6-12.0)		
	SGR	ICp	36.0 (24.9-41.7)	19.3 (16.8-21.0)	50.6 (47.9-53.0)	23.2 (20.0-27.7)	56.3 (54.1-58.3)	32.5 (29.7-35.1)	Harmful	<5/5
		Weibull	38.8 (23.5-53.1)	18.4 (4.6-28.1)	46.9 (31.3-59.9)	26.1 (8.6-36.5)	53.5 (38.1-64.5)	34.2 (12.8-44.5)		
		Polynomial Regression	42.2 (40.0-44.4)	17.3 (15.3-19.5)	51.5 (48.9-54.1)	23.0 (20.0-26.3)	57.5 (55.6-59.3)	29.1 (25.8-32.5)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
4-CP	Yield	ICp	32.5 (30.5-34.2)	20.6 (14.7-23.3)	29.9 (26.7-33.3)	14.7 (8.8-26.7)	29.3 (26.4-31.7)	15.2 (9.3-20.3)	Harmful	5/10
		Weibull	32.2 (23.6-40.0)	19.9 (7.2-26.0)	28.4 (19.9-38.1)	15.2 (5.0-21.2)	28.1 (20.0-36.6)	15.6 (5.7-21.3)		
		Polynomial Regression	32.9 (29.5-37.3)	14.7 (12.0-18.0)	28.8 (25.6-31.8)	11.7 (9.5-14.3)	27.9 (24.4-31.6)	11.9 (9.5-14.8)		
	SGR	ICp	42.2 (38.3-46.5)	24.5 (22.7-26.3)	43.6 (42.6-44.8)	24.2 (21.2-27.5)	44.9 (41.6-47.1)	24.9 (22.6-26.9)	Harmful	5/10
		Weibull	40.7 (32.2-53.8)	25.4 (10.5-32.1)	41.3 (32.5-53.7)	25.4 (11.5-32.3)	42.3 (34.0-55.1)	26.7 (12.6-33.4)		
		Polynomial Regression	45.8 (43.1-48.5)	21.4 (18.6-24.5)	46.4 (45.0-48.0)	21.5 (19.5-23.7)	47.8 (45.4-50.0)	23.0 (20.3-26.0)		
2,3-DCP	Yield	ICp	11.2 (6.3-14.2)	4.9 (0.1-11.4)	7.9 (5.8-10.0)	2.0 (1.3-4.4)	6.3 (4.9-7.5)	2.1 (0.9-3.8)	Toxic	<2/2
		Weibull	10.5 (6.8-14.9)	4.9 (1.3-7.3)	7.0 (3.7-11.2)	2.5 (0.3-4.5)	6.0 (3.0-9.7)	2.2 (0.2-3.9)		
		Polynomial Regression	10.7 (7.6-14.6)	4.1 (1.8-7.4)	7.5 (6.6-8.7)	2.5 (1.8-3.3)	6.7 (5.9-7.7)	2.2 (1.4-3.0)		
	SGR	ICp	14.6 (12.7-16.4)	7.7 (3.4-11.6)	13.1 (12.3-13.9)	5.9 (3.5-7.9)	13.3 (12.5-14.1)	5.5 (4.2-6.6)	Harmful	<2/2
		Weibull	14.2 (10.3-19.0)	7.8 (2.8-10.6)	12.5 (8.4-17.4)	6.1 (1.9-8.9)	12.4 (8.3-17.6)	5.9 (1.7-8.7)		
		Polynomial Regression	15.4 (12.5-18.4)	6.4 (4.0-9.6)	13.2 (12.1-14.4)	5.1 (4.4-6.0)	13.3 (12.3-14.4)	5.1 (4.4-6.0)		
2,4-DCP	Yield	ICp	6.6 (6.1-7.2)	4.2 (3.8-4.7)	5.7 (4.7-6.6)	3.1 (1.9-4.4)	5.2 (3.9-6.1)	2.4 (0.7-3.5)	Toxic	<2/2
		Weibull	6.6 (4.8-8.3)	4.1 (1.3-5.4)	5.6 (4.0-7.3)	3.2 (1.1-4.3)	5.0 (3.1-7.0)	2.4 (0.5-3.6)		
		Polynomial Regression	7.3 (6.7-8.0)	3.2 (2.7-3.7)	6.2 (5.5-7.1)	2.5 (1.8-3.4)	5.7 (5.0-6.4)	2.0 (1.3-2.8)		
	SGR	ICp	10.0 (6.4-11.9)	5.4 (4.9-5.9)	9.3 (8.2-10.4)	4.9 (4.1-5.7)	9.3 (8.3-10.3)	4.7 (3.7-5.4)	Toxic	2/4
		Weibull	10.5 (7.1-14.4)	5.2 (1.6-7.5)	9.5 (6.5-12.9)	4.8 (1.6-6.9)	9.1 (6.2-12.4)	4.6 (1.5-6.6)		
		Polynomial Regression	10.7 (9.6-11.9)	4.6 (4.0-5.2)	9.6 (8.8-10.4)	4.1 (3.6-4.7)	9.1 (8.6-9.8)	3.8 (3.4-4.3)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
2,5-DCP	Yield	ICp	10.1 (7.1-13.1)	4.4 (0.1-7.2)	7.2 (6.3-8.2)	2.0 (1.3-5.4)	6.1 (4.4-7.1)	1.7 (0.9-3.1)	Toxic	<2/2
		Weibull	9.6 (5.7-14.6)	3.9 (0.8-6.3)	6.9 (3.6-10.9)	2.5 (0.3-4.4)	5.6 (2.6-9.1)	1.9 (0.1-3.5)		
		Polynomial Regression	10.1 (8.1-12.6)	3.6 (2.2-5.5)	7.4 (6.7-8.3)	2.5 (1.9-3.1)	6.4 (5.6-7.3)	1.9 (1.2-2.7)		
	SGR	ICp	14.9 (12.8-17.7)	6.5 (4.6-9.0)	13.1 (12.3-14.2)	5.7 (4.7-6.7)	12.5 (11.6-13.3)	5.2 (2.7-6.5)	Harmful	<2/2
		Weibull	13.9 (9.6-19.1)	7.0 (2.3-10.0)	12.3 (8.2-17.5)	5.9 (1.7-8.7)	11.7 (7.7-16.8)	5.5 (1.5-8.2)		
		Polynomial Regression	15.2 (13.1-17.5)	6.0 (4.5-8.1)	13.2 (12.2-14.2)	5.1 (4.5-5.8)	12.4 (11.4-13.4)	4.7 (4.1-5.5)		
2,6-DCP	Yield	ICp	12.9 (11.4-14.5)	2.9 (2.4-3.8)	9.3 (8.0-10.4)	3.1 (2.3-4.1)*	8.5 (6.7-10.3)	3.1 (2.2-4.9)	Toxic	<5/5
		Weibull	11.1 (2.8-19.4)	3.0 (0.1-6.8)	9.1 (2.4-14.9)	2.9 (0.1-5.9)	8.5 (2.4-13.5)	2.9 (0.1-5.7)		
		Polynomial Regression	14.0 (12.9-15.1)	3.1 (2.3-4.0)	11.8 (10.9-12.9)	2.5 (1.5-3.4)	11.2 (10.2-12.3)	2.3 (1.3-3.4)		
	SGR	ICp	24.4 (21.3-27.2)	8.0 (5.0-10.7)	21.6 (19.6-23.6)	7.9 (6.0-9.3)	21.5 (20.5-22.5)	7.9 (6.6-9.4)	Harmful	<5/5
		Weibull	22.3 (12.7-34.6)	8.7 (1.5-14.4)	20.0 (12.1-29.6)	8.5 (1.8-13.5)	18.9 (13.0-26.0)	9.5 (3.1-13.6)		
		Polynomial Regression	23.5 (22.0-25.1)	8.1 (7.3-9.0)	20.5 (19.5-21.6)	7.7 (7.0-8.4)	18.3 (17.5-19.2)	7.5 (6.9-8.1)		
3,4-DCP	Yield	ICp	4.0 (3.6-4.5)	2.2 (1.4-3.0)	3.7 (3.2-4.1)	1.7 (0.9-2.5)	3.8 (3.3-4.2)	1.6 (0.7-3.2)	Toxic	0.663/1.325
		Weibull	3.9 (2.8-5.2)	2.2 (0.7-3.0)	3.5 (2.4-4.6)	1.8 (0.6-2.6)	3.5 (2.5-4.6)	1.9 (0.7-2.6)		
		Polynomial Regression	4.0 (3.4-4.6)	1.7 (1.2-2.2)	3.5 (3.0-4.0)	1.4 (1.1-1.9)	3.5 (3.0-4.1)	1.5 (1.1-1.9)		
	SGR	ICp	5.5 (5.0-6.1)	3.2 (2.7-3.6)	5.4 (5.3-5.7)	3.1 (2.8-3.5)	5.5 (4.9-6.0)	3.3 (3.0-3.6)	Toxic	0.663/1.325
		Weibull	5.4 (4.3-7.1)	3.4 (1.3-4.2)	5.3 (4.1-7.0)	3.3 (1.3-4.2)	5.4 (4.3-7.3)	3.4 (1.2-4.3)		
		Polynomial Regression	6.0 (5.6-6.5)	2.9 (2.4-3.4)	5.9 (5.7-6.2)	2.8 (2.4-3.1)	6.1 (5.6-6.5)	2.9 (2.4-3.4)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
3,5-DCP	Yield	ICp	2.5 (2.3-2.6)	1.3 (0.9-1.8)	2.2 (1.9-2.4)	1.1 (0.7-1.4)	2.2 (1.9-2.5)	1.1 (0.8-1.4)	Toxic	0.8/1.6
		Weibull	2.5 (1.8-3.2)	1.5 (0.5-2.0)	2.1 (1.5-2.8)	1.2 (0.4-1.6)	2.2 (1.6-2.8)	1.3 (0.5-1.7)		
		Polynomial Regression	2.6 (2.4-2.9)	1.2 (1.0-1.4)	2.2 (1.9-2.4)	1.0 (0.7-1.2)	2.2 (1.9-2.6)	1.0 (0.8-1.4)		
	SGR	ICp	3.5 (3.4-3.6)	2.0 (1.8-2.2)	3.4 (3.3-3.5)	1.8 (1.6-2.1)	3.5 (3.3-3.6)	1.6 (1.9-1.7)	Toxic	0.8/1.6
		Weibull	3.4 (2.7-4.6)	2.2 (0.9-2.7)	3.3 (2.6-4.3)	2.0 (0.9-2.6)	3.3 (2.7-4.5)	2.1 (0.9-2.7)		
		Polynomial Regression	3.8 (3.7-3.9)	1.9 (1.7-2.1)	3.7 (3.5-3.8)	1.8 (1.6-2.0)	3.8 (3.6-4.0)	1.9 (1.6-2.1)		
2,3,4-TCP	Yield	ICp	1.5 (0.6-2.3)	0.5 (0.1-1.7)	1.4 (0.8-2.0)	0.5 (0.1-1.3)	1.6 (1.2-1.9)	0.4 (0.2-0.9)	Toxic	<0.5/0.5
		Weibull	1.5 (0.5-2.7)	0.4 (0.1-0.9)	1.5 (0.5-2.6)	0.4 (0.1-0.9)	1.6 (0.6-2.8)	0.4 (0.1-0.9)		
		Polynomial Regression	1.8 (1.3-2.7)	0.5 (0.1-1.2)	1.7 (1.4-2.3)	0.5 (0.2-0.9)	1.8 (1.5-2.2)	0.5 (0.3-0.8)		
	SGR	ICp	4.1 (2.5-5.1)	1.1 (0.3-1.8)	4.3 (3.8-4.9)	1.1 (0.6-1.6)	4.5 (4.2-4.8)	1.3 (1.0-1.7)	Toxic	<0.5/0.5
		Weibull	3.2 (1.8-5.2)	1.1 (0.2-2.0)	3.4 (1.9-5.5)	1.2 (0.2-2.1)	3.7 (2.2-5.6)	1.5 (0.3-2.4)		
		Polynomial Regression	3.5 (2.8-4.4)	1.2 (0.7-1.9)	3.7 (3.3-4.2)	1.3 (0.9-1.7)	4.1 (3.8-4.5)	1.5 (1.2-1.8)		
2,3,5-TCP	Yield	ICp	1.4 (0.5-2.4)	0.5 (0.1-1.2)	1.4 (0.5-2.3)	0.4 (0.2-1.2)	1.4 (0.7-2.1)	0.4 (0.2-1.2)	Toxic	<0.6/0.6
		Weibull	1.4 (0.7-2.2)	0.5 (0.1-0.9)	1.2 (0.5-2.0)	0.4 (0.1-0.8)	1.2 (0.5-1.9)	0.4 (0.1-0.8)		
		Polynomial Regression	1.5 (1.1-2.3)	0.5 (0.3-1.0)	1.4 (1.0-2.2)	0.4 (0.2-0.9)	1.3 (1.0-1.8)	0.4 (0.3-0.7)		
	SGR	ICp	2.7 (2.3-3.1)	1.0 (0.2-1.6)	2.7 (2.3-3.1)	1.2 (0.4-1.9)	2.7 (2.4-3.0)	1.3 (0.5-1.8)	Toxic	<0.6/0.6
		Weibull	2.4 (1.8-3.2)	1.3 (0.5-1.8)	2.4 (1.8-3.2)	1.4 (0.6-1.8)	2.4 (1.8-3.2)	1.4 (0.6-1.9)		
		Polynomial Regression	2.7 (2.3-3.1)	1.2 (0.8-1.7)	2.7 (2.3-3.1)	1.2 (0.8-1.7)	2.7 (2.4-3.0)	1.2 (0.9-1.6)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
2,3,6-TCP	Yield	ICp	5.1 (4.8-5.4)	1.8 (1.6-2.3)	4.6 (4.1-5.0)	1.7 (1.2-2.2)	3.4 (3.2-3.7)	1.6 (1.3-1.9)	Toxic	<1/1
		Weibull	4.6 (3.1-6.3)	2.3 (0.7-3.3)	4.2 (2.8-5.8)	2.0 (0.6-3.0)	3.4 (2.2-4.7)	1.6 (0.4-2.4)		
		Polynomial Regression	4.5 (4.3-4.8)	1.8 (1.7-2.0)	4.2 (3.8-4.5)	1.7 (1.4-2.1)	3.5 (3.3-3.7)	1.4 (1.2-1.6)		
	SGR	ICp	6.1 (4.9-7.1)	2.3 (0.8-3.9)	6.6 (6.2-6.9)	4.1 (3.7-4.4)	6.1 (5.8-6.3)	3.1 (2.8-3.3)	Toxic	1/2
		Weibull	5.7 (3.8-8.2)	2.7 (0.7-4.0)	6.5 (4.7-8.4)	3.9 (1.3-5.1)	5.9 (4.1-7.8)	3.2 (1.1-4.4)		
		Polynomial Regression	6.8 (6.4-7.3)	2.9 (2.7-3.2)	6.7 (6.3-7.2)	2.9 (2.6-3.2)	6.0 (5.7-6.3)	2.6 (2.4-2.7)		
2,4,5-TCP	Yield	ICp	2.8 (2.3-3.2)	1.2 (0.7-1.5)	2.3 (1.6-3.0)	1.1 (0.4-1.6)	2.3 (1.7-2.9)	1.1 (0.4-1.5)	Toxic	<0.5/0.5
		Weibull	2.8 (1.7-4.1)	1.2 (0.3-1.8)	2.4 (1.5-3.5)	1.0 (0.2-1.6)	2.3 (1.4-3.3)	1.0 (0.2-1.6)		
		Polynomial Regression	2.8 (2.5-3.2)	1.1 (0.8-1.4)	2.4 (2.1-2.8)	0.9 (0.7-1.3)	2.3 (2.0-2.6)	0.9 (0.7-1.2)		
	SGR	ICp	4.7 (4.4-5.0)	1.9 (1.6-2.3)	4.4 (4.2-4.6)	1.9 (1.5-2.3)	4.2 (3.9-4.4)	2.1 (1.6-2.4)	Toxic	0.5/1
		Weibull	4.3 (2.9-6.0)	2.1 (0.7-3.1)	4.1 (2.8-5.7)	2.0 (0.7-2.9)	4.1 (2.8-5.5)	2.1 (0.8-2.9)		
		Polynomial Regression	4.7 (4.4-5.0)	2.0 (1.7-2.3)	4.5 (4.2-4.8)	1.9 (1.6-2.2)	4.4 (4.1-4.7)	1.9 (1.6-2.1)		
2,4,6-TCP	Yield	ICp	3.8 (3.3-4.4)	1.2 (0.4-3.0)	3.2 (2.1-4.1)	0.8 (0.2-2.8)	2.9 (2.3-3.5)	0.7 (0.3-2.2)	Toxic	<1/1
		Weibull	3.7 (2.1-5.8)	1.5 (0.2-2.4)	3.0 (1.4-4.8)	1.0 (0.1-1.9)	2.6 (1.2-4.2)	0.9 (0.1-1.7)		
		Polynomial Regression	3.9 (3.5-4.4)	1.4 (1.1-1.7)	3.3 (2.7-4.1)	1.0 (0.4-1.8)	3.0 (2.6-3.6)	0.9 (0.5-1.4)		
	SGR	ICp	6.4 (5.9-6.8)	2.9 (2.5-3.4)	6.1 (5.5-6.8)	2.6 (1.8-3.5)	5.8 (5.2-6.4)	2.6 (2.1-3.0)	Toxic	<1/1
		Weibull	6.1 (4.1-8.6)	2.9 (0.9-4.3)	5.7 (3.7-8.2)	2.6 (0.7-4.0)	5.5 (3.5-8.0)	2.5 (0.6-3.8)		
		Polynomial Regression	6.4 (6.0-6.9)	2.5 (2.2-2.9)	6.0 (5.4-6.8)	2.3 (1.8-2.9)	5.8 (5.3-6.3)	2.2 (1.9-2.6)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
3,4,5-TCP	Yield	ICp	0.8 (0.1-1.4)	0.07 (0.03-0.2)	0.6 (0.1-1.0)	0.06 (0.03-0.14)	0.5 (0.1-0.8)	0.05 (0.04-0.08)	Very toxic	<0.1/0.1
		Weibull	0.6 (0.1-1.1)	0.1 (0.01-0.3)	0.4 (0.1-0.9)	0.1 (0.01-0.2)	0.3 (0.1-0.7)	0.05 (0.01-0.2)		
		Polynomial Regression	0.8 (0.6-1.0)	0.1 (0.01-0.3)	0.7 (0.5-0.8)	0.1 (0.01-0.3)	0.6 (0.5-0.7)	0.1 (0.01-0.2)		
	SGR	ICp	1.5 (1.3-1.6)	0.3 (0.01-0.8)	1.4 (1.3-1.5)	0.5 (0.01-0.8)	1.3 (1.2-1.4)	0.5 (0.08-0.7)	Toxic	<0.1/0.1
		Weibull	1.3 (0.8-2.0)	0.6 (0.09-0.9)	1.3 (0.8-1.8)	0.6 (0.1-0.9)	1.2 (0.8-1.7)	0.6 (0.1-0.9)		
		Polynomial Regression	1.4 (1.3-1.6)	0.4 (0.3-0.6)	1.3 (1.2-1.4)	0.4 (0.3-0.6)	1.2 (1.1-1.4)	0.4 (0.3-0.5)		
2,3,4,5-TeCP	Yield	ICp	0.9 (0.8-1.0)	0.3 (0.1-0.4)	0.6 (0.2-0.8)	0.2 (0.1-0.3)	0.4 (0.3-0.5)	0.14 (0.12-0.17)	Very toxic	<0.335/0.335
		Weibull	0.8 (0.4-1.1)	0.4 (0.1-0.6)	0.5 (0.1-0.8)	0.2 (0.01-0.3)	0.4 (0.01-0.6)	0.1 (0.01-0.3)		
		Polynomial Regression	0.8 (0.7-0.9)	0.3 (0.2-0.4)	0.6 (0.5-0.7)	0.2 (0.1-0.3)	0.5 (0.4-0.6)	0.1 (0.09-0.2)		
	SGR	ICp	1.15 (1.09-1.2)	0.5 (0.3-0.7)	1.1 (1.0-1.2)	0.4 (0.2-0.6)	1.05 (1.0-1.1)	0.3 (0.2-0.4)	Toxic	<0.335/0.335
		Weibull	1.1 (0.8-1.3)	0.6 (0.3-0.9)	1.0 (0.6-1.3)	0.5 (0.1-0.7)	0.9 (0.5-1.3)	0.4 (0.08-0.7)		
		Polynomial Regression	1.0 (0.9-1.1)	0.4 (0.2-0.5)	1.0 (0.9-1.1)	0.4 (0.3-0.5)	0.9 (0.8-1.0)	0.3 (0.2-0.4)		
2,3,4,6-TeCP	Yield	ICp	6.0 (5.2-6.6)	2.7 (0.9-4.0)	6.1 (4.6-7.0)	2.6 (1.0-6.2)	5.8 (4.9-6.5)	2.5 (0.4-4.0)	Toxic	<1.25/1.25
		Weibull	5.7 (3.2-7.6)	2.9 (0.4-4.3)	5.7 (3.5-7.3)	3.2 (0.7-4.6)	5.4 (3.2-7.0)	2.9 (0.6-4.2)		
		Polynomial Regression	6.1 (5.2-6.9)	2.6 (1.8-3.6)	6.1 (4.9-7.1)	2.8 (1.7-4.3)	5.8 (4.8-6.7)	2.5 (1.7-3.7)		
	SGR	ICp	8.2 (7.8-8.8)	4.9 (3.6-5.6)	8.2 (7.9-8.8)	5.5 (4.0-6.4)	8.3 (7.9-8.7)	5.5 (4.8-6.1)	Toxic	2.5/5
		Weibull	8.0 (5.7-10.9)	4.9 (0.7-6.5)	8.2 (6.2-10.0)	5.6 (1.6-6.9)	8.2 (6.2-10.0)	5.6 (1.7-6.9)		
		Polynomial Regression	8.3 (7.8-8.8)	4.8 (4.2-5.5)	8.4 (7.9-8.8)	5.3 (4.6-6.1)	8.4 (7.9-8.8)	5.4 (4.8-6.0)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
2,3,5,6-TeCP	Yield	ICp	6.2 (4.9-7.2)	2.9 (1.6-4.0)	6.1 (5.0-6.8)	3.2 (0.9-5.0)	5.3 (4.1-6.2)	2.3 (1.0-3.5)	Toxic	1/2
		Weibull	6.0 (4.1-8.4)	3.0 (0.9-4.3)	5.9 (4.2-7.8)	3.3 (1.1-4.5)	5.0 (3.3-6.9)	2.4 (0.7-3.6)		
		Polynomial Regression	6.3 (5.2-7.5)	2.5 (1.7-3.6)	6.1 (4.8-7.6)	2.6 (1.7-3.9)	5.0 (4.2-6.0)	2.0 (1.5-2.8)		
	SGR	ICp	8.7 (8.0-9.2)	4.5 (3.7-5.3)	8.6 (8.1-8.9)	4.9 (3.9-6.7)	8.6 (7.8-9.3)	4.7 (3.8-5.3)	Toxic	2/4
		Weibull	8.2 (6.4-10.7)	5.0 (2.2-6.4)	8.3 (6.7-11.1)	5.3 (2.1-6.6)	8.2 (6.4-10.7)	5.0 (2.2-6.4)		
		Polynomial Regression	9.7 (9.0-10.2)	4.7 (3.9-5.6)	9.7 (8.9-10.4)	4.9 (3.9-6.0)	9.2 (8.5-9.8)	4.3 (3.6-5.1)		
PCP	Yield	ICp	6.1 (1.9-8.2)	2.2 (1.5-2.7)	5.1 (2.7-6.8)	1.6 (0.9-2.2)	5.2 (4.5-5.7)	1.6 (1.0-1.8)	Toxic	1/2
		Weibull	5.7 (3.4-8.7)	2.3 (0.5-3.7)	4.7 (2.7-7.2)	1.8 (0.3-3.0)	4.6 (2.8-7.0)	1.9 (0.4-3.1)		
		Polynomial Regression	6.2 (5.2-7.4)	2.3 (1.6-3.3)	4.9 (3.9-6.3)	1.8 (1.1-2.9)	4.8 (4.2-5.4)	1.8 (1.3-2.5)		
	SGR	ICp	10.2 (9.7-10.7)	3.6 (2.8-5.0)	9.5 (9.0-9.9)	4.5 (2.2-6.0)	9.4 (9.1-9.7)	4.8 (4.3-5.2)	Toxic	1/2
		Weibull	9.1 (7.2-11.4)	5.6 (3.0-7.1)	8.6 (6.7-11.0)	5.2 (2.6-6.7)	8.7 (6.9-11.1)	5.4 (2.7-6.8)		
		Polynomial Regression	10.1 (9.4-10.7)	5.0 (4.1-6.0)	9.6 (8.9-10.3)	4.6 (3.6-5.7)	9.7 (9.4-10.0)	4.7 (4.2-5.2)		
Pyrogallol	Yield	ICp	4.6 (4.0-5.2)	1.8 (0.9-3.0)	3.6 (3.1-4.2)	1.4 (0.7-2.2)	3.5 (2.7-4.7)	1.3 (0.5-2.3)	Toxic	<1/1
		Weibull	4.2 (2.7-6.1)	1.9 (0.5-2.9)	3.5 (2.1-5.3)	1.5 (0.3-2.4)	3.4 (1.9-5.1)	1.4 (0.2-2.2)		
		Polynomial Regression	4.3 (3.8-4.8)	1.7 (1.3-2.1)	3.7 (3.3-4.2)	1.4 (1.0-1.8)	3.6 (3.1-4.3)	1.3 (0.9-1.8)		
	SGR	ICp	6.5 (6.0-6.9)	3.3 (2.6-4.1)	6.4 (6.0-6.7)	2.9 (2.3-3.5)	6.6 (6.1-7.2)	3.0 (2.3-3.9)	Toxic	<1/1
		Weibull	6.3 (4.4-8.5)	3.4 (1.1-4.7)	6.1 (4.2-8.5)	3.1 (0.9-4.4)	6.3 (4.3-8.7)	3.2 (1.0-4.6)		
		Polynomial Regression	6.6 (6.0-7.2)	2.7 (2.3-3.2)	6.4 (6.0-6.9)	2.6 (2.3-2.9)	6.7 (6.0-7.5)	2.7 (2.2-3.2)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
Hq	Yield	ICp	85.6 (54.7-103.3)	26.8 (8.0-39.4)	72.0 (50.8-98.5)	17.6 (10.5-33.8)	69.9 (50.1-94.4)	18.0 (0.1-32.1)	Harmful	10/20
		Weibull	72.0 (48.8-100.3)	35.1 (11.3-51.0)	60.1 (35.9-91.5)	24.3 (5.0-39.2)	59.8 (36.1-90.4)	24.4 (5.2-39.2)		
		Polynomial Regression	81.1 (69.7-92.0)	33.7 (17.8-48.9)	74.3 (64.5-84.0)	24.8 (11.9-37.5)	74.0 (62.5-85.1)	24.7 (8.6-40.3)		
	SGR	ICp	105.9 (96.7-113.4)	53.4 (25.3-92.9)	104.4 (97.2-112.4)	54.7 (35.6-80.7)	105.7 (99.0-112.2)	60.3 (42.7-83.5)	Not toxic	10/20
		Weibull	98.7 (85.0-120.1)	69.0 (49.9-80.8)	96.7 (82.2-118.6)	66.1 (45.2-78.5)	98.7 (85.4-121.0)	69.7 (50.9-81.0)		
		Polynomial Regression	109.4 (103.2-115.1)	62.3 (47.6-75.7)	106.3 (100.4-128.8)	58.1 (45.8-80.8)	114.0 (113.3-115.2)	65.0 (54.0-74.4)		
ClHq	Yield	ICp	6.2 (4.6-7.9)	1.8 (0.1-4.4)	5.7 (3.0-8.7)	1.6 (0.1-4.4)	3.8 (2.8-6.0)	1.3 (0.3-2.4)	Toxic	<1/1
		Weibull	5.9 (3.0-10.0)	1.9 (0.2-3.5)	5.3 (2.5-9.4)	1.6 (0.1-3.1)	4.1 (1.8-7.1)	1.2 (0.1-2.4)		
		Polynomial Regression	6.0 (5.0-7.2)	2.0 (1.2-3.0)	5.5 (4.0-7.8)	1.8 (0.7-3.5)	4.3 (3.5-5.4)	1.4 (0.9-2.1)		
	SGR	ICp	11.6 (10.3-13.1)	4.3 (2.8-6.3)	11.8 (10.5-13.4)	4.2 (2.3-6.7)	10.7 (9.6-11.7)	3.4 (2.4-5.1)	Harmful	<1/1
		Weibull	11.0 (7.0-19.3)	4.4 (0.7-7.0)	11.1 (6.9-20.4)	4.3 (0.6-6.9)	10.1 (6.0-19.3)	3.6 (0.4-6.0)		
		Polynomial Regression	11.0 (10.1-12.1)	4.1 (3.3-5.0)	11.6 (9.9-13.0)	4.4 (2.9-6.3)	10.2 (9.0-11.4)	3.6 (2.9-4.5)		
TetraClHq	Yield	ICp	5.7 (4.7-6.6)	1.6 (0.6-2.6)	5.2 (4.5-5.9)	2.1 (0.6-3.0)	4.2 (3.4-5.2)	1.4 (0.4-2.5)	Toxic	<1/1
		Weibull	5.1 (2.9-7.9)	2.0 (0.4-3.3)	4.8 (3.0-7.1)	2.1 (0.5-3.3)	4.0 (2.2-6.3)	1.5 (0.2-2.5)		
		Polynomial Regression	5.5 (4.6-6.5)	1.9 (1.3-2.6)	5.0 (4.4-5.6)	1.8 (1.4-2.3)	4.2 (3.8-4.8)	1.4 (1.1-1.9)		
	SGR	ICp	8.8 (8.0-9.7)	4.1 (2.7-5.0)	8.5 (8.0-9.1)	4.3 (3.6-4.9)	8.5 (8.0-9.0)	3.8 (3.1-4.5)	Toxic	<1/1
		Weibull	8.0 (5.9-10.5)	4.5 (1.9-6.0)	8.0 (6.1-10.4)	4.6 (2.0-6.1)	7.7 (5.6-10.3)	4.2 (1.7-5.7)		
		Polynomial Regression	9.0 (8.2-9.7)	3.9 (3.2-4.8)	8.9 (8.4-9.4)	4.0 (3.5-4.5)	8.6 (8.2-9.1)	3.7 (3.2-4.2)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
Cat	Yield	ICp	30.3 (23.3-36.7)	9.6 (6.4-14.3)	17.7 (16.1-19.7)	7.0 (5.2-10.2)	15.9 (13.5-18.2)	5.7 (4.3-7.8)	Harmful	<10/10
		Weibull	30.1 (10.9-52.9)	8.5 (0.2-17.8)	20.4 (3.0-37.3)	5.0 (0.1-12.3)	16.7 (0.6-31.4)	3.7 (0.1-10.3)		
		Polynomial Regression	35.7 (31.7-38.9)	9.5 (6.7-12.7)	27.4 (25.2-30.0)	5.1 (3.1-7.3)	24.3 (21.9-27.0)	3.1 (0.8-5.5)		
	SGR	ICp	66.5 (60.0-73.2)	17.7 (14.2-21.3)	60.8 (58.0-63.4)	14.9 (12.8-16.6)	59.7 (55.6-63.4)	14.1 (11.5-16.6)	Harmful	<10/10
		Weibull	57.2 (33.6-88.0)	22.5 (4.4-37.0)	51.4 (28.0-82.4)	18.5 (2.7-32.1)	50.1 (26.7-81.3)	17.6 (2.3-31.0)		
		Polynomial Regression	61.2 (54.9-69.2)	21.3 (18.1-25.3)	57.2 (53.7-59.2)	18.3 (16.2-20.1)	55.5 (52.2-59.1)	17.4 (15.1-19.9)		
4-ClCat	Yield	ICp	7.0 (5.2-8.3)	3.4 (3.1-3.7)	6.0 (5.2-7.4)	2.9 (2.3-3.5)	5.4 (4.3-6.8)	2.2 (1.0-3.0)	Toxic	<1.5/1.5
		Weibull	7.0 (4.6-9.8)	3.3 (1.0-4.9)	6.1 (4.0-8.7)	2.8 (0.7-4.3)	5.3 (3.1-7.9)	2.2 (0.4-3.5)		
		Polynomial Regression	7.1 (6.5-7.6)	2.9 (2.6-3.3)	6.3 (5.8-6.8)	2.5 (2.1-3.0)	5.6 (4.9-6.3)	2.1 (1.6-2.6)		
	SGR	ICp	10.8 (10.2-11.2)	5.3 (4.4-6.3)	10.7 (10.2-11.0)	5.2 (4.4-6.1)	10.6 (10.0-11.0)	5.0 (4.1-6.5)	Harmful	1.5/3.0
		Weibull	10.5 (7.5-14.1)	5.7 (2.0-7.8)	10.3 (7.4-13.9)	5.5 (1.9-7.6)	10.1 (7.1-13.8)	5.3 (1.8-7.4)		
		Polynomial Regression	11.3 (10.6-11.9)	4.8 (4.4-5.3)	11.1 (10.5-11.6)	4.7 (4.2-5.1)	10.9 (10.2-11.6)	4.5 (4.0-5.0)		
3,5-DiClCat	Yield	ICp	1.3 (1.2-1.5)	0.5 (0.2-0.8)	1.0 (0.8-1.4)	0.4 (0.2-0.7)	1.0 (0.7-1.3)	0.4 (0.2-0.7)	Toxic	<0.5/0.5
		Weibull	1.2 (0.7-1.8)	0.5 (0.1-0.8)	1.0 (0.6-1.5)	0.4 (0.1-0.7)	1.0 (0.6-1.5)	0.4 (0.1-0.6)		
		Polynomial Regression	1.2 (1.1-1.4)	0.5 (0.4-0.6)	1.1 (0.9-1.2)	0.4 (0.3-0.5)	1.1 (0.9-1.2)	0.4 (0.3-0.5)		
	SGR	ICp	2.0 (1.9-2.1)	1.0 (0.8-1.1)	1.9 (1.8-2.0)	0.9 (0.6-1.2)	2.0 (1.8-2.2)	0.9 (0.7-1.2)	Toxic	<0.5/0.5
		Weibull	1.9 (1.3-2.6)	1.0 (0.3-1.4)	1.8 (1.2-2.5)	0.9 (0.3-1.3)	1.9 (1.3-2.6)	1.0 (0.3-1.4)		
		Polynomial Regression	2.0 (1.9-2.1)	0.8 (0.7-0.9)	1.9 (1.7-2.1)	0.8 (0.7-0.9)	2.0 (1.8-2.2)	0.8 (0.7-0.9)		

Table 4.10. Continued

Compound	Reponse Variable	Method	48h		72h		96h		Toxic Class	NOEC/LOEC
			IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)	IC_{50} (mg/L)	IC_{20} (mg/L)		
Res	Yield	ICp	282.9 (255.0-306.5)	90.0 (50.3-153.0)	200.0 (120.0-260.6)	69.9 (34.3-99.5)	177.7 (117.7-244.7)	66.0 (32.8-94.5)	Not toxic	<50/50
		Weibull	252.9 (163.1-362.5)	116.5 (30.3-175.5)	180.6 (101.4-277.3)	71.0 (10.9-118.1)	167.8 (93.1-257.8)	65.8 (9.5-110.0)		
		Polynomial Regression	335.3 (307.1-363.2)	99.1 (61.2-135.9)	285.4 (243.7-327.9)	41.5 (8.5-102.3)	272.4 (229.9-317.9)	26.5 (4.4-88.9)		
	SGR	ICp	377.4 (334.1-414.6)	224.8 (198.6-252.8)	348.1 (315.0-381.6)	160.5 (104.1-238.5)	342.5 (308.9-375.3)	159.0 (101.7-228.9)	Not toxic	50/100
		Weibull	372.2 (279.7-492.5)	217.4 (80.3-286.5)	332.1 (231.1-456.2)	170.2 (56.0-240.9)	327.5 (227.4-450.3)	167.6 (54.5-237.4)		
		Polynomial Regression	423.2 (368.9-624.4)	178.5 (147.8-240.1)	354.9 (315.9-397.9)	143.7 (118.0-176.1)	352.8 (313.2-397.2)	141.9 (117.8-172.7)		
4-ClRes	Yield	ICp	62.3 (52.9-70.5)	34.1 (8.6-51.2)	54.7 (47.0-60.9)	23.4 (12.7-34.1)	53.3 (47.1-57.9)	25.2 (13.7-32.6)	Harmful	10/20
		Weibull	60.6 (43.4-78.8)	34.4 (12.0-46.5)	52.4 (36.8-69.3)	28.2 (10.1-39.1)	50.7 (36.1-66.3)	27.8 (10.8-38.2)		
		Polynomial Regression	60.4 (48.6-76.5)	26.3 (17.4-39.3)	52.2 (42.2-62.3)	22.9 (16.2-31.2)	49.3 (42.2-55.3)	21.9 (17.1-27.4)		
	SGR	ICp	82.5 (70.3-94.0)	48.8 (39.6-55.6)	78.6 (68.4-90.7)	45.7 (39.5-52.6)	78.5 (69.5-88.2)	46.3 (41.9-49.9)	Harmful	20/40
		Weibull	80.6 (64.1-108.6)	51.2 (18.3-64.3)	77.8 (60.2-103.0)	47.6 (17.5-61.1)	77.8 (60.4-103.4)	48.0 (17.0-61.4)		
		Polynomial Regression	85.6 (80.3-91.2)	39.9 (33.1-46.9)	84.2 (79.7-88.8)	38.8 (33.3-44.3)	84.3 (81.1-87.7)	39.0 (35.1-42.9)		
4,6-DiClRes	Yield	ICp	15.0 (13.7-16.0)	10.5 (6.9-12.1)	13.9 (12.1-15.2)	8.7 (4.7-12.0)	13.4 (12.3-14.1)	7.9 (4.8-9.5)	Harmful	5/10
		Weibull	15.3 (12.1-17.5)	11.0 (6.5-13.4)	13.5 (11.0-15.9)	9.1 (5.9-11.2)	13.1 (10.7-15.5)	8.7 (5.8-10.7)		
		Polynomial Regression	15.1 (12.6-17.9)	7.3 (5.2-10.0)	12.9 (10.8-15.5)	6.1 (4.4-8.3)	12.5 (11.1-14.0)	6.1 (4.84-7.5)		
	SGR	ICp	17.5 (16.6-18.3)	12.1 (11.0-13.0)	17.0 (16.2-17.7)	11.7 (10.6-12.5)	16.9 (16.3-17.5)	11.5 (10.8-12.0)	Harmful	5/10
		Weibull	18.1 (0.5-20.7)	13.4 (0.1-16.5)	17.4 (11.5-20.1)	12.3 (2.7-15.3)	17.3 (11.6-20.1)	12.2 (2.9-15.2)		
		Polynomial Regression	19.8 (17.9-21.9)	9.5 (7.8-11.6)	18.6 (16.8-20.3)	8.7 (7.3-10.2)	18.5 (17.4-19.6)	8.7 (7.7-9.7)		

[a]:confidence intervals; [b]: toxic class based on 72h ICP calculations; [c]: NOEC and LOEC values were calculated using 72h toxicity data

C. vulgaris revealed dose-dependent responses to chemicals tested in this study. The response of algae to phenol is provided as an example in Figure 4.9.

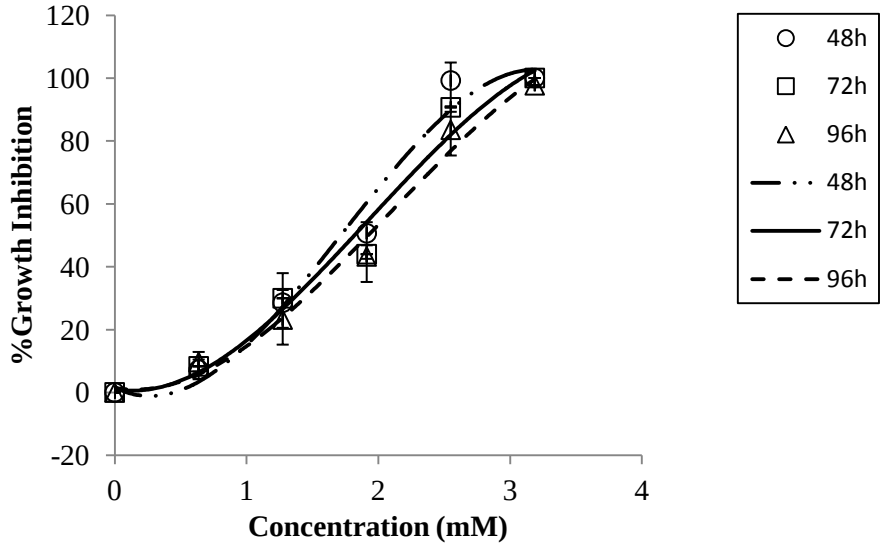


Figure 4.9. Growth response of *Chlorella vulgaris* to phenol (bars represent experimental errors)

Based upon average specific growth rate, the IC_{50} and associated confidence intervals for 48 h and 96 h was found to overlap, which suggests that the toxicity of the tested phenols to *C. vulgaris* did not change significantly between these durations (Table 4.10).

Based on the IC_{50} values, the least toxic compound was found to be phenol, while the most toxic compound was 2,3,4,5-tetrachlorophenol regardless of exposure duration or response variable (Table 4.10). As expected, either the IC_{50} or IC_{20} values based upon average specific growth rate were found to be higher than those based upon yield due to the mathematical basis of the respective approaches (OECD, 2006). Similar to the case of *D. tertiolecta*, the IC_{50} values did not change significantly based on the calculation method as the confidence intervals were found to overlap.

According to toxicity classification provided in Table 4.10, based on IC_{50} calculations following the yield method, two chemicals were classified as not toxic, seven compounds as harmful, 19 chemicals as toxic and two chemicals as very toxic. When the IC_{50} calculated by the specific growth rate method (SGR) were used for classification, the toxic class of one chemical (hydroquinone) was reduced from harmful to not toxic. Likewise, the

toxic class of two chemicals (3,4,5-trichlorophenol and 2,3,4,5-tetrachlorophenol) was reduced from very toxic to toxic and five chemicals (2,3-dichlorophenol, 2,5-dichlorophenol, 2,6-dichlorophenol, chlorohydroquinone and 4-chlorocatechol) were classified as harmful instead of toxic (Table 4.10).

In a ring test conducted by the participation of 18 laboratories, the algal toxicity of 3,5-DCP to freshwater alga *P. subcapitata* was found to be 3.4±1.30 mg/L (ISO, 2004). In this study, the toxicity of 3,5-DCP to another freshwater alga *C. vulgaris* was found to be 3.3±0.2 mg/L; therefore, it can be said that the results obtained in this study concur with international standards for algal toxicity testing. Additionally, the toxicity of 3,5-DCP, the reference toxicant, was determined twice for freshwater alga *C. vulgaris* and it was found that the 95% confidence intervals overlapped for the two IC_{50} values suggesting that the algal response did not change significantly during the course of the study.

It was observed that the *ortho* substituted chlorophenols were less toxic than *meta* or *para* substituted congeners to *C. vulgaris*, similar to the previous findings on the toxicity of chlorophenols to marine alga *D. tertiolecta.* Furthermore, the simultaneous occurrence of chlorine atoms at both *ortho* positions (i.e., 2- and 6- positions relative to the OH moiety) resulted in reduced toxicity of these congeners compared to those with equal number of chlorine atoms. For example, 2,6-DCP was found to be the least toxic dichlorophenol. Likewise, 2,3,6-TCP and 2,4,6-TCP were less toxic than other trichlorophenols; whereas 3,4,5-TCP, which has both *ortho* positions unoccupied, was the most toxic trichlorophenol (Table 4.9). This is probably because, as stated by Escher et al. (1996), heterodimer formation is hindered due to localization of the charge as well as steric interactions by the simultaneous occupation of both *ortho* positions in a chlorophenol molecule which in turn reduces its toxicity.

4.2.1. Correlation of *C. vulgaris* toxicity with hydrophobicity

The relationship between the toxicity of compounds acting by polar narcosis should be strongly related to their hydrophobicity as modelled by Log *P* (Vighi et al., 2009). Analysis of the data set presented in this study revealed that there is, as expected, a strong trend between the toxicity of polar narcotics and hydrophobicity as described by Log *P*:

pT_{96h}= 0.84 Log *P* – 1.42 (4.3)

(*n*=19, r^2=0.93, r^2_{cv}=0.91, *RMSE*=0.21)

In the European Union Technical Guidance Document (EU TGD, 2003), several models employing the hydrophobicity parameter, Log *P*, were proposed for the use of these relationships in preliminary aquatic risk assessment. It was underlined in the technical guidance document that only the experimental data that were generated according to OECD test guidelines or comparable methods were used in the reported models. Since the toxicological assays in this study were conducted according to OECD algal growth inhibition guidelines (OECD, 2006) and standard methods (APHA-AWWA-WEF, 1998), the experimental data generated in the present study meets the quality criteria implied in the document. Therefore, the equation developed in this study for polar narcotics (Eq. 4.3) can be included in preliminary risk assessment scenarios for *C. vulgaris*. It should be noted that the equations reported are valid in the Log *P* range 1.46 - 4.01.

4.2.2. QSTRs for freshwater alga *C. vulgaris*

The Kolmogorov-Smirnov test revealed that toxicity of phenols to *C. vulgaris* followed a normal distribution (p>0.05). The distribution of the toxicity data (pT_{96}) is illustrated in Figure 4.10.

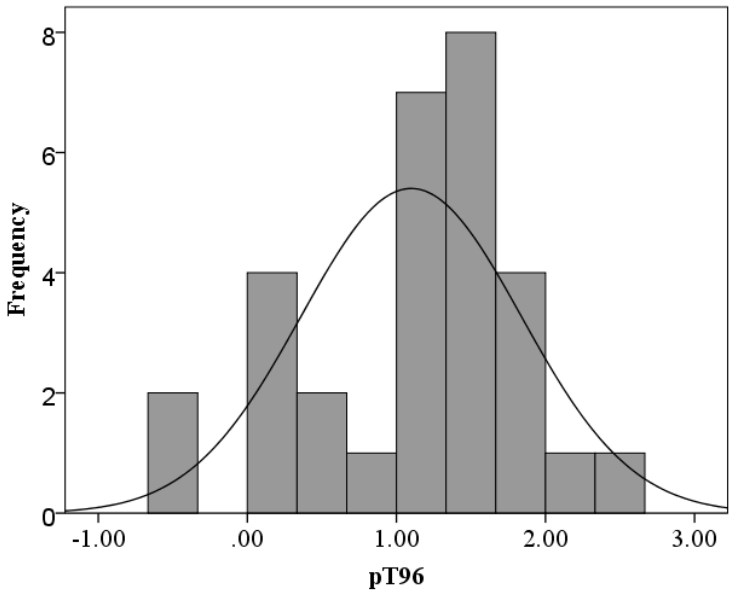
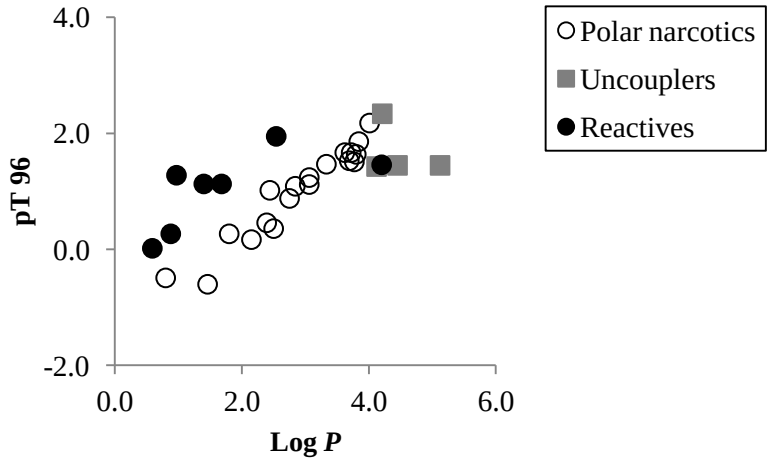
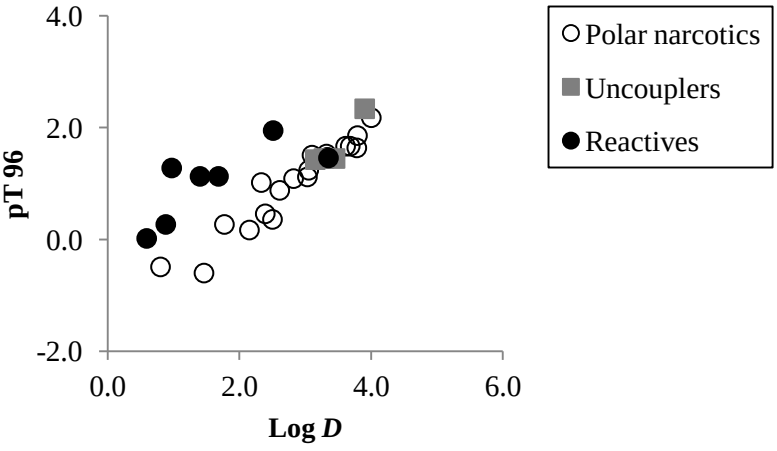


Figure 4.10. Histogram of the 96 h toxicity data for *C. vulgaris* (pT_{96})

Toxicity data to *C.* v*ulgaris* (pT_{96}) for 30 phenols are presented in Table 4.11 together with the expected MOA, interaction type and descriptor values for each compound. The compounds exhibited a wide range of algal toxicity from -0.60 to 2.34 [on a log(mM) scale]. The relationship between the algal toxicity of phenols and hydrophobicity as described by Log *P* and Log *D* is illustrated in Figure 4.11(a) and Figure 4.11(b), respectively.



(a)



(b)

Figure 4.11. Relationship between the algal toxicity and hydrophobicity of phenols as described by (a) Log *P* and (b) Log *D*

Table 4.11. CAS number, algal toxicity of phenols to *C. vulgaris*, expected mode of action (MOA) and descriptor values

Compound ID	CAS No.	Compound	Abbreviation	*Expected MOA*	pT_{96} (mM)	Log P	*p*Ka	Log D[c]	$E_{(HOMO-LUMO)}$ (eV)	Tm	nHBonds	AEI
Training set												
1	108-95-2	Phenol	P	Polar narcosis	-0.60	1.46[a]	9.99[a]	1.46	9.47	3.05	0	-
2	95-57-8	2-chlorophenol	2-CP	Polar narcosis	0.17	2.15[a]	8.56[a]	2.15	9.07	4.61	0	-
3	108-43-0	3-chlorophenol	3-CP	Polar narcosis	0.36	2.50[a]	9.12[a]	2.50	9.25	5.02	0	-
5	576-24-9	2,3-dichlorophenol	2,3-DCP	Polar narcosis	1.09	2.84[a]	7.70[a]	2.82	8.93	5.78	0	-
6	120-83-2	2,4-dichlorophenol	2,4-DCP	Polar narcosis	1.24	3.06[a]	7.89[a]	3.05	8.82	5.99	0	-
8	87-65-0	2,6-dichlorophenol	2,6-DCP	Polar narcosis	0.88	2.75[a]	6.79[a]	2.61	8.95	5.46	0	-
9	95-77-2	3,4-dichlorophenol	3,4-DCP	Polar narcosis	1.47	3.33[a]	8.63[a]	3.33	8.84	6.20	0	-
10	591-35-5	3,5-dichlorophenol	3,5-DCP	Polar narcosis	1.67	3.62[a]	8.18[a]	3.61	9.15	6.09	0	-
11	15950-66-0	2,3,4-trichlorophenol	2,3,4-TCP	Polar narcosis	1.64	3.80[a]	7.59[a]	3.78	8.68	6.67	0	-
12	933-78-8	2,3,5-trichlorophenol	2,3,5-TCP	Polar narcosis	1.86	3.84[a]	7.23[a]	3.79	8.65	6.62	0	-
13	933-75-5	2,3,6-trichlorophenol	2,3,6-TCP	Polar narcosis	1.51	3.77[a]	5.80[a]	3.10	8.80	6.36	0	-
14	95-95-4	2,4,5-trichlorophenol	2,4,5-TCP	Polar narcosis	1.67	3.72[a]	7.40[a]	3.68	8.56	6.71	0	-
16	609-19-8	3,4,5-trichlorophenol	3,4,5-TCP	Polar narcosis	2.18	4.01[a]	7.84[a]	4.00	8.72	6.84	0	-
17	4901-51-3	2,3,4,5-tetrachlorophenol	2,3,4,5-TCP	Uncoupling	2.34	4.21[a]	6.35[a]	3.90	8.47	7.21	0	-
19	935-95-5	2,3,5,6-tetrachlorophenol	2,3,5,6-TCP	Uncoupling	1.43	4.12[a]	5.44[a]	3.15	8.51	6.96	0	-
20	87-86-5	Pentachlorophenol	PCP	Uncoupling	1.45	5.12[a]	4.70[a]	3.45	8.35	7.46	0	-
21	87-66-1	1,2,3-trihydroxybenzene	Pyrogallol	Electrophilic	1.28	0.97[b]	9.01[a]	0.97	9.23	4.03	2	12.15[d]
23	615-67-8	Chlorohydroquinone	ClHq	Electrophilic	1.13	1.40[a]	8.81[a]	1.40	8.74	5.20	0	12.69[d]
24	87-87-6	Tetrachlorohydroquinone	TetraClHq	Electrophilic	1.46	4.20[b]	5.58[b]	3.35	8.20	7.07	0	11.02[d]
25	120-80-9	Catechol	Cat	Electrophilic	0.27	0.88 [a]	9.45[a]	0.88	9.16	3.52	1	13.05[d]
26	2138-22-9	4-chlorocatechol	4ClCat	Electrophilic	1.13	1.68[b]	8.67[b]	1.68	8.88	5.41	1	12.02[d]
28	108-46-3	Resorcinol	Res	Polar narcosis	-0.49	0.80[a]	9.32[a]	0.80	9.33	3.76	0	-
29	95-88-5	4-chlororesorcinol	4ClRes	Polar narcosis	0.27	1.80[a]	7.58[b]	1.77	8.98	5.23	0	-
30	137-19-9	4,6-dichlororesorcinol	46DiClres	Polar narcosis	1.02	2.44[b]	6.89[b]	2.33	8.69	6.27	0	-
Test set												
4	106-48-9	4-chlorophenol	4-CP	Polar narcosis	0.46	2.39[a]	9.41[a]	2.39	9.06	5.15	0	-
7	583-78-8	2,5-dichlorophenol	2,5-DCP	Polar narcosis	1.12	3.06[a]	7.51[a]	3.03	8.77	5.82	0	-
15	88-06-2	2,4,6-trichlorophenol	2,4,6-TCP	Polar narcosis	1.53	3.69[a]	6.23[a]	3.32	8.68	6.50	0	-
18	58-90-2	2,3,4,6-tetrachlorophenol	2,3,4,6-TCP	Uncoupling	1.45	4.45[a]	5.22[a]	3.28	8.51	7.07	0	-
22	123-31-9	Hydroquinone	Hq	Electrophilic	0.02	0.59[a]	10.90[a]	0.59	8.92	3.90	0	14.11[d]
27	13673-92-2	3,5-dichlorocatechol	35DiClCat	Electrophilic	1.95	2.54[b]	7.57[b]	2.51	8.77	6.25	1	11.12[e]

[a]: Experimental value; [b]: Estimated value; [c]: Log D was calculated at the test *p*H of 6.36 using Eq. (3.1); [d]: taken from Aptula et al. (2005b); [e]: the AEI value of 3,5-DiClCat was estimated using equation 3 provided in Aptula et al. (2005b) and toxicity of this compound to *T. pyriformis* as determined by Bajot et al. (2011).

Log *D* was found to be a better parameter to quantify hydrophobicity than Log *P*, demonstrating the importance of ionization on partitioning for the tested phenols. Examination of Figure 4.11a revealed that some compounds were less toxic than predicted by the polar narcosis model (Eq. 4.3) for *C. vulgaris*. These compounds were noted to be three uncouplers (2,3,4,6-tetrachlorophenol, 2,3,5,6-tetrachlorophenol and pentachlorophenol) and a reactive toxicant, namely, tetrachlorohydroquinone (TetraClHq). The algal toxicity of polar narcotics and uncouplers significantly improved when Log *D*, instead of Log *P*, was used to quantify hydrophobicity (Figure 4.11b). Accordingly, the regression analysis revealed a strong correlation between algal toxicity (pT_{96}) and hydrophobicity (Log *D*) for both the polar narcotics and uncouplers:

$$pT_{96} = 0.90 \text{ Log } D - 1.50 \quad (4.4)$$

(n=23, r^2=0.92, r^2_{cv}=0.90, *RMSE*=0.21)

Figure 4.11(b) also revealed that phenols with two or more hydroxy groups positioned on the ring in *ortho* or *para* configuration to each other (as represented by black dots), which are potentially capable of being oxidized to reactive species (Bajot et al., 2011), displayed algal toxicity in excess of that predicted by hydrophobicity as described by Log *D*. Interestingly, the toxicity of tetrachlorohydroquinone (TetraClHq) had a reasonable agreement with Log *D* and lies well on the fit line with the rest of polar narcotics (residual=+0.06, Figure 4.11(b)). As discussed earlier, this is probably because with increasing hydrophobicity the toxicity of electrophiles and pro-electrophiles converges on a narcosis mechanism in terms of both the observed physiological effects and the fit to the narcosis model (Aptula et al., 2005b).

The relationship between the algal toxicity and hydrophobicity implicitly suggested that another parameter describing the electrophilicity of the chemicals was required to model the toxicity of all the tested phenols towards *C. vulgaris*. From an organic chemistry perspective, the reactivity *per se* should be modeled by a reaction rate constant or activation energy (Aptula et al., 2009). One such descriptor is the Activation Energy Index (AEI), which can be calculated from the computed highest occupied molecular orbital (HOMO) and HOMO-1 orbital energies of the electrophiles and the intermediates based on the principle that the activation energy for the reaction of an electrophile with a

nucleophile can be modeled by the energy differences between the occupied molecular orbitals of the metastable intermediate and the electrophile. Detailed information on the chemical basis of the parameter as well as an illustrative example for its calculation can be found in Aptula et al. (2005c).

The AEI parameter has been shown to correlate well with the aquatic toxicity of several polyphenols to *Tetrahymena pyriformis*, with some exceptions like methyl-substituted hydroquinones (Aptula et al., 2005b). Since the data set presented in this study contains chlorine as the sole substituent in the *ortho* or *para* substituted polyphenols, the AEI parameter should provide a suitable basis to evaluate the reactivity of these phenols towards *C. vulgaris*. The AEI values of the reactive phenols and the adjusted toxicity data following the rules derived by Aptula et al. (2005b), are provided in Table 4.12.

Table 4.12. The unadjusted and adjusted algal toxicity of reactive phenols

Compound	$pT_{96\ (mmol/l)}$	$pT_{96(adj)\ (mmol/l)}$
pyrogallol	1.28	1.28
hydroquinone	0.02	-0.58
chlorohydroquinone	1.13	1.13
tetrachlorohydroquinone	1.46	0.86
catechol	0.27	-0.03
4-chlorocatechol	1.13	1.13
3,5-dichlorocatechol	1.95	1.65

As expected, a significant relationship was obtained between the AEI parameter and algal toxicity of reactive phenols (Eq. 4.5):

$$pT_{96} = -0.57\ AEI + 8.05 \qquad (4.5)$$

(n=7, r^2=0.85, r^2_{cv}=0.73, *RMSE*=0.24)

It should be indicated that Aptula et al. (2005b) adjusted the toxicity data to take into account the contribution of alternative electrophilic centers to toxicity which might result from the oxidation of the parent compound. Following the rules derived by Aptula et al. (2005b), the algal toxicity of the reactive phenols was adjusted (pT_{96adj}, Table 4.12) and an equally good relationship was observed between the adjusted algal toxicity (pT_{96adj}) and AEI, only following the removal of TetraClHq (Eq. 4.6):

pT_{96adj} = -0.79 AEI + 10.69 (4.6)

(n=6, r^2=0.87, r^2_{cv}=0.75, *RMSE*=0.29)

It was interesting to note that the hydrophobicity model employing Log *D* (Eq. 4.4) provided a much better prediction for the toxicity of TetraClHq (residual +0.06) than any of the reactivity relationships developed (residuals by Eq. 4.5 and Eq. 4.6 were +0.31 and +0.49, respectively). The reason why the reactivity of TetraClHq is not well-predicted by the AEI parameter as compared to Log *D* is, as mentioned earlier, probably due to the dominant polar narcotic mechanism of the compound, which in turn might have suppressed the contribution of reactive mechanisms to its toxicity. Aptula and Roberts (2006) stated that despite being reactive some compounds act by a narcosis mechanism because their narcotic toxicity is greater than their reactive toxicity. It seems likely that TetraClHq is one such chemical which exhibits algal toxicity mainly by polar narcotic mechanisms. Another reason behind the suppressed reactivity of TeClHq might be the low bioavailability of this compound at the test *p*H as it has a lower *p*Ka compared to other reactive toxicants tested in this study (Table 4.11) assuming that the unionized form of the molecule is more bioavailable, hence more toxic, than the ionized form (Ertürk and Saçan, 2012 – references therein). In conclusion, the analysis of the mechanistic makeup of the data set revealed two general mechanisms of action: one hydrophobicity-dependent and one reactivity-dependent.

4.2.3. Global models

Efforts to construct global QSTRs using the novel algal toxicity data of phenols towards *C. vulgaris* resulted initially in a large pool of models with acceptable training set statistics. For practical purposes, those displaying unacceptable test set performance was eliminated from the model pool. Finally, the remaining models were further tested in their external predictivity on the external data set obtained from Aruoja et al. (2011). Among the remaining global QSTRs, only those displaying acceptable external validation performance are listed in Table 4.13. The Log *P* model reported by Aruoja et al. (2011) for the external validation set compounds was also reported in Table 4.13. The descriptor values of the external data set compounds are provided in Appendix D (Table D3).

Table 4.13. Statistical overview of the algal QSTRs for *C. vulgaris*

Model 1[a]														
pT_{96} = 0.49 Log *D* - 0.88 $E_{(HOMO-LUMO)}$ + 7.58														
Training set (*n*=24)			**Test set (*n*=6)**						**External validation set (*n*=58)**					
r^2	r^2_{cv}	*RMSE*	R^2	R_0^2	$R_0'^2$	*k*	*k'*	$RMSE_{te}$	R^2	R_0^2	$R_0'^2$	*k*	*k'*	$RMSE_{val}$
0.82	0.73	0.33	0.64	0.54	0.63	0.91	0.99	0.40	0.66	0.52	0.66	0.85	1.06	0.35
Model 2[a]														
pT_{96} = 0.58 Tm - 2.28														
Training set (*n*=24)			**Test set (*n*=6)**						**External validation set (*n*=58)**					
r^2	r^2_{cv}	*RMSE*	R^2	R_0^2	$R_0'^2$	*k*	*k'*	$RMSE_{te}$	R^2	R_0^2	$R_0'^2$	*k*	*k'*	$RMSE_{val}$
0.81	0.77	0.33	0.78	0.75	0.78	0.94	1.00	0.31	0.82	0.82	0.76	0.83	1.12	0.32
Model 3[b]														
pT_{96} = 0.62Tm +0.61nHBonds – 2.58														
Training set (*n*=24)			**Test set (*n*=6)**						**External validation set (*n*=58)**					
r^2	r^2_{cv}	*RMSE*	R^2	R_0^2	$R_0'^2$	*k*	*k'*	$RMSE_{te}$	R^2	R_0^2	$R_0'^2$	*k*	*k'*	$RMSE_{val}$
0.84	0.79	0.30	0.94	0.94	0.93	1.03	0.95	0.18	0.85	0.82	0.72	0.79	1.15	0.35
Model 4[c]														
pIC_{50} = 0.65 Log *P* - 0.71														
a: The models were built following the removal of pyrogallol (1,2,3-trihydroxybenzene) from the training set (n=23) which was a response outlier. **b**: Two compounds (aniline and 3,4-dimethylaniline) were found to be response outliers in the external set (n=56). **c**: from Aruoja et al. (2011)									**External validation set (*n*=58)**					
									R^2	R_0^2	$R_0'^2$	*k*	*k'*	$RMSE_{val}$
									0.60	0.60	0.40	1.00	0.88	-

4.2.4. Mechanistic interpretation of the global models

Global QSTRs offer the possibility of producing a single model that may cover a number of different mechanisms of toxic action. For regulatory usage, it is also crucial to have an insight as to how the global model statistics relate to the predictions of the individual mechanisms of toxic action within the model's applicability domain. Such an understanding enables confidence to be better assigned to the resulting predictions when the models are used in a real life scenario (Enoch et al., 2008). Therefore, it is also of our interest to investigate the levels of confidence that can be assigned for each mechanism of toxic action present in the global QSTRs. To this end, the errors associated with the prediction of each global QSTR, as judged by the *RMSE* values, were calculated separately for polar narcotics, uncouplers and reactive compounds in the data set and compared to those of local models (Eq. 4.4 for polar narcotics and uncouplers, Eq. 4.5 for reactive chemicals) which served as a comparative basis to evaluate the performance of global model statistics (Table 4.14).

Table 4.14. *RMSE* values for each MOA in the data set

Model ID	***RMSE***		
	Polar narcotics (*n*=19)	**Respiratory uncouplers (*n*=4)**	**Reactive compounds (*n*=7)**
Model 1	0.26	0.32	0.71
Model 2	0.26	0.45	0.60
Model 3	0.24	0.44	0.26
Eq. 4.4	0.22	0.19	-
Eq. 4.5	-	-	0.24

Model 1, employing the *p*H corrected hydrophobicity Log *D* and $E_{(HOMO\text{-}LUMO)}$, the latter describing the energy difference between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), can be regarded as a response-surface model where the parameters quantify the bio-uptake and electrophilicity of the compounds, respectively. From a theoretical perspective, the energy gap between HOMO and LUMO of a molecule has been regarded as an important stability index (Karelson et al., 1996). Consequently, its presence in the two-descriptor model is believed to account for the reactivity of the electrophilic phenols. Negative coefficient of this descriptor in model 1 (Table 4.13) indicates that toxicity decreases with increasing HOMO-LUMO gap, which is consistent with the explanations of Karelson et al. (1996)

who stated that a large gap between HOMO and LUMO implies high stability for the molecule in the sense of its lower reactivity in chemical reactions.

The response-surface approach can be considered as an attempt to model a variety of non-congeneric chemicals acting via non-specific interactions without covalent bond rearrangements. In this approach, the response of an organism to the presence of a toxicant in the environment is considered as a continuum of effects brought about by the uptake of the chemical into the biophase and its interaction with the site of action (Dimitrov et al., 2003). In this respect, the response-surface model developed in this study is easily interpretable from a toxicological perspective. However, it was found to be the least predictive among the global QSTRs. Examination of the RMSE values (Table 4.14) indicated that the compounds capable of being oxidized to quinones were likely to have toxicity in excess of the response-surface. In the response-surface approach, the chemicals not included in the surface are believed to possess well-defined reactive groups, which could bind by covalent bond rearrangements (Dimitrov et al., 2003). Previously, Cronin et al. (2002) analyzed the toxicity of phenols towards *Tetrahymena pyriformis* using the response-surface approach (employing Log *D* and E_{LUMO}) and the researchers also observed disappointing results for phenols transformable to their electrophilic forms (i.e., reactive phenols). On the other hand, the model error, compared to that of Eq. 4.4, was not too high for compounds exhibiting toxicity via non-covalent interactions (i.e., polar narcosis and respiratory uncoupling). For the external data set predictions, the response-surface model lacked the predictive power and the model explained only 6% more variance in toxicity as compared to the Log *P* model reported by Aruoja et al. (2011, Table 4.13). For the response-surface model developed in this study, detailed analysis revealed that Log *D*, as calculated in the test *p*H for the external set (*p*H=8.0, Aruoja et al. (2011)), produced a much worse prediction (n=30, r^2=0.38) than Log *P* (n=30, r^2=0.85) for phenols. A separate response-surface model (Eq. 4.7) using Log *P*, instead of Log *D*, to quantify hydrophobicity and $E_{(HOMO-LUMO)}$ to quantify electrophilicity of the external set compounds was found to be more predictive than model 1 (Table 4.13):

$$p\mathrm{EC}_{50} = 0.50\ \mathrm{Log}\ P - 0.88\ E_{(HOMO-LUMO)} + 7.56 \qquad (4.7)$$

(n=58, r^2=0.74, r^2_{cv}=0.71, *RMSE*=0.31)

The response-surface approach employing either Log *P* or Log *D* for the measure of hydrophobicity has high interpretability and portability which has been successfully employed for predicting aquatic toxicity (Cronin et al., 2002; Dimitrov et al., 2003; Enoch et al., 2008). The analysis conducted in this study revealed that toxicity data generated in different algal test systems may vary in their dependence on the use of Log *P* or Log *D* to explain the bio-uptake of chemicals. This might be due to the distinct characteristics of algal test systems such as the test *p*H, test organism, chemical composition of the growth medium, illumination level, temperature, etc.

Two other global QSTRs were also obtained at the end of the modeling process (Table 4.13). Model 2 is comprised of only one descriptor, namely Tm (T total size index / weighted by atomic masses), which was obtained from Dragon 5.4 software. It is a WHIM (Weighted Holistic Invariant Molecular) descriptor where the statistical indices are calculated on the atoms projected onto the 3 principal components obtained from weighted covariance matrices of the atomic coordinates (Todeschini and Consonni, 2009). WHIM descriptors have been successfully used in many QSAR studies for several classes of compounds and different endpoints (Todeschini and Gramatica, 1997 – and references therein). It was interesting to note that the correlation between Log *D* and Tm of the tested phenols was found to be relatively high (r^2=0.79), suggesting that Tm carries significant information regarding the hydrophobicity of the molecules. In addition, considering the ability of Tm in explaining aniline toxicity in the external set (n=28, r^2=0.80), it is likely that this descriptor encodes some mechanistic information other than hydrophobicity which alone was not able to provide a reliable prediction for anilines (Aruoja et al., 2011). Therefore, for simple aromatic compounds that do not have a good correlation to either Log *P* or Log *D*, Tm might be a promising descriptor to predict algal toxicity. On the other hand, model 2 employing Tm as the sole descriptor produced relatively high *RMSE* for reactive phenols (*RMSE*=0.60, Table 4.14). Consequently, the use of Tm to predict algal toxicity arising due to the covalent interactions between the electrophiles and nucleophiles is not recommended.

Model 3 as compared to model 2, includes additionally the descriptor nHBonds, which quantifies the number of intra-molecular hydrogen bonds in a compound (Todeschini and Consonni, 2000). The additional information carried by this descriptor

improved both the training and test set performance of model 3 as compared to model 2, resulting in reliable overall toxicity predictions for phenols to *C. vulgaris*. It was also noted that error of the model in predicting the toxicity for reactive toxicants was the lowest among the global models (Table 4.14). Inspection of nHBonds values revealed that only 1,2,3-trihydroxybenzene (pyrogallol) and catechols had values for this descriptor while the remaining chemicals, including the hydroquinones, had zero values (Table 4.11). This implies that the intra-molecular hydrogen bonds are defined, according to nHBonds, only for polyphenols which have *ortho* OH substitution. From an organic chemistry perspective, ortho-dihydroxybenzenes (i.e., catechols) can be oxidized abiotically and enzymatically to *ortho*-quinones, which can act as Michael acceptor electrophiles. *Ortho*-quinones are much more reactive as Michael acceptors than the isomeric *para*-quinones (i.e., hydroquinones) because the carbon-oxygen-carbon-oxygen group (COCO) is more electronegative than the carbon-oxygen group (Aptula et al., 2005b). In line with the greater activating effect of the COCO group, catechol was found to be more toxic than hydroquinone to *C. vulgaris*. As for pyrogallol, this compound should be more toxic than catechol because the quinone from of 1,2,3-trihydroxybenzene should be more electrophilic than the *ortho* quinone from catechol (Aptula et al., 2005b). Accordingly, pyrogallol was found to be more toxic to *C. vulgaris* than catechol. It is logical to assume that the presence of nHBonds in model 3 seems to correct for the potency order observed among the polyhydroxybenzenes. It seems that nHBonds acts as an indicator variable to discriminate reactivity of catechols and pyrogallol over other compounds.

The concept of reactivity is an important issue in predictive toxicology (Aptula et al., 2005b; Schultz et al., 2006; Schwöbel et al., 2011). Briefly, reactivity includes a spectrum of conjugation, substitution, and addition reactions during the interaction of electrophiles with nucleophilic sites in peptides, proteins, and nucleic acids (Schultz et al., 2006; Schwöbel et al., 2011). However, these interactions between a reactive compound and cellular components are not specific as they can disrupt many different cellular and/or organismal processes (Schultz et al., 2006). From a predictive perspective, representation of these processes has resulted in means of grouping compounds into categories for the particular end point being investigated based on mechanistic similarity (Schwöbel et al., 2011). From this point of view, the number of intra-molecular hydrogen bonds in model 3 (Table 4.13) offers a basic grouping of phenols with two or more hydroxyl groups in *ortho*

positions (i.e., catechols, pyrogallol) indicating their excess toxicity in the *C. vulgaris* test system.

As for the external predictivity of model 3, the inclusion of nHBonds did not significantly improve the model statistics and the external prediction resulted in two response outliers, namely, aniline (ID=V16, Table D3, in Appendix D) and 3,4-dimethylaniline (ID=V37, Table D3, in Appendix D). Yet, considering the external validation set predictions of model 2 and model 3 together, both models were significantly more predictive than the Log *P* model reported by Aruoja et al (2011).

4.2.5. Applicability domain of the global models

The applicability domain of each model was constructed using the plot of standardized residuals against leverages (Figure 4.12).

It was found that all the compounds in the test and external validation sets were within the applicability domain of each global model as the leverage value of each compound was lower than the critical hat value (h*) of each model, with the exception of pyrogallol (ID=21) for model 3, which was a good leverage compound (Figure 4.12). The unique nature of the compound, having three OH in its structure, is probably the reason why it was a structural outlier. The predicted toxicity by each global model as well as the calculated leverages and standardized residuals can be found in the supplementary materials to this study (Appendix D, Table D4).

Pyrogallol was detected as a response outlier in model 1 and model 2 (Table 4.13). As mentioned earlier, the inability of these models to explain reactivity in the data set is probably the reason of the poor prediction. Model 3, with the inclusion of an indicator variable, nHBonds, was able to offer a more reliable prediction for this compound.

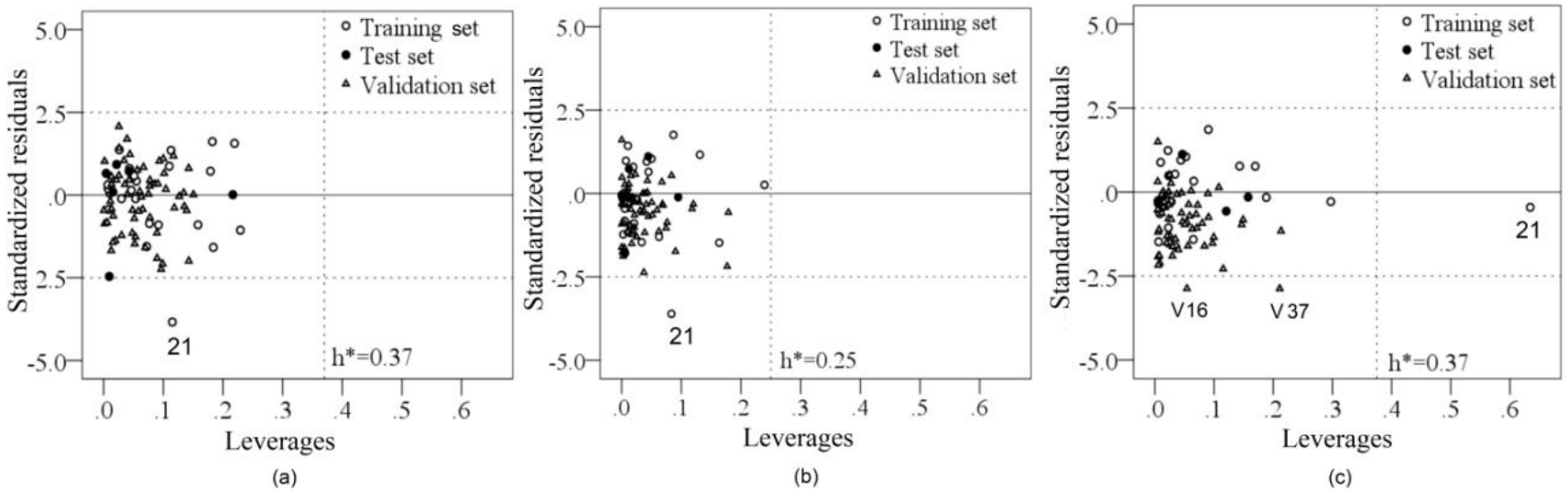


Figure 4.12. The Williams plot of the algal QSTRs developed for *C. vulgaris*

4.2.6. Counter Propagation Artificial Neural Networks

With the introduction of Tm and nHBonds descriptors as input variables, the two-parameter CP ANN modeling was performed employing three different network architectures (5x5, 6x6 and 7x7) to cover all investigated objects as well as a range of learning iterations (100, 200, 400, 800, 1000 iterations). The modeling was performed on the training set (24 compounds), employing the leave-one-out cross-validation procedure for optimization of the technical parameters. The best CP ANN model was selected according to the highest r^2_{cv} value which was obtained from a network of 6x6 dimensions and 800 epochs.

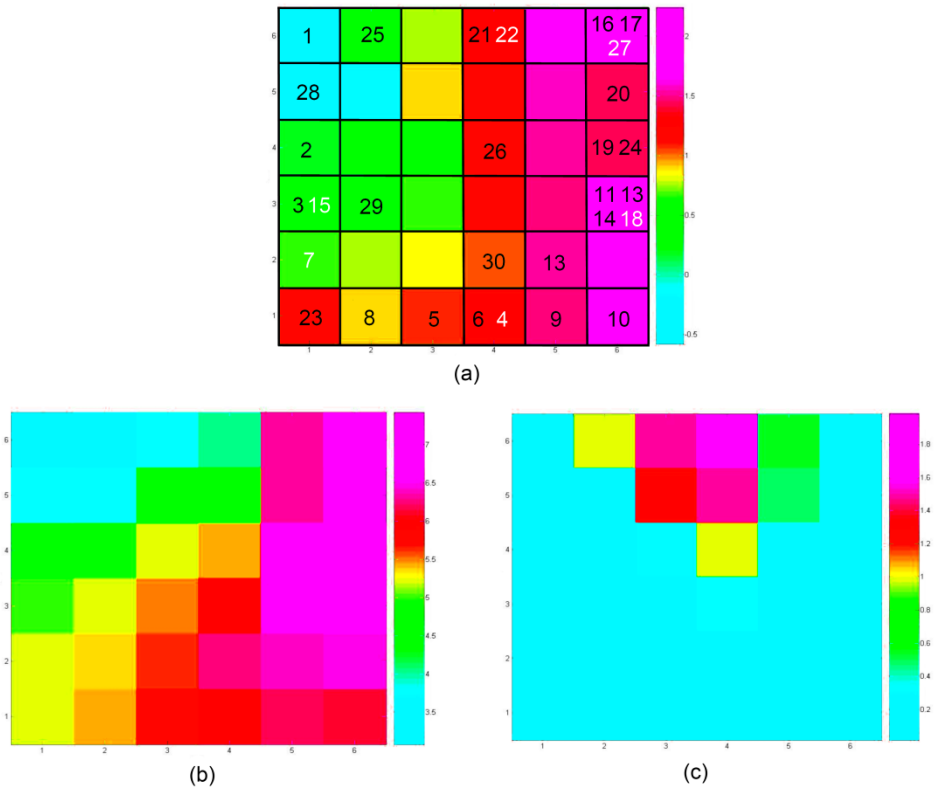


Figure 4.13. Output layers of the two parameter CP ANN model: a) toxicity (pT_{96}) b) Tm and c) nHBonds (Compound ID in black and white refers to training and test set compounds, respectively).

The visual analysis of the output layers for the obtained CP ANN model (Figure 4.13) highlighted the good agreement between the weight maps of toxicity (Figure 4.13a) and Tm (Figure 4.13b). On the other hand, the weight map of the indicator variable, nHBonds

(Figure 4.13c), follow a different color distribution compared to the biological toxicity map (Figure 4.13a).

The statistical performance of the CP ANN model and its comparison with the linear model (model 3 in Table 4.13) is provided in Table 4.15.

Table 4.15. Statistical performance of CP ANN and MLR models for *Chlorella vulgaris*

Model	**Training set**			**Test set**		
	n	r^2	r^2_{cv}	*n*	Q^2_{F3}	*RMSE*
CP ANN	24	0.94	0.84	6	0.95	0.18
MLR	24	0.84	0.79	6	0.94	0.18

The results revealed that the while the training set performance of the two-parameter CP ANN model was statistically superior over that of the MLR model, the test set performance of both models were very close (Table 4.15). Since the real usefulness of a QSAR lies in its ability to predict the toxicity of untested chemicals, it can be argued that both models were equally capable of predicting the toxicity of untested phenols; as such, they can be treated as having similar performance in terms of predictivity.

4.3. Interspecies toxicity correlations

The use of toxicity data on one species to predict the toxicity of the same set of chemicals to another species is a promising area which has received little attention (Zvinavashe et al., 2008). One of the aims of this thesis was to investigate if freshwater toxicity data can be used to predict marine toxicity. In this respect, the models that can predict the marine algal toxicity using toxicity data on freshwater alga *C. vulgaris* were investigated. To this end, several molecular descriptors, if necessary, were used to improve the raw correlations between the organisms in an interpretable manner. The interspecies correlation between freshwater and marine algae was also evaluated based on the mechanism of action of each tested phenols. Finally, the interspecies toxicity correlations were further extended to include other aquatic species including freshwater algal species *P. subcapitata*, the marine bacterium *V. fischeri*, and other freshwater organisms such as the protozoan *T. pyriformis*, daphnid *D. magna* and fish *P. promelas*.

4.3.1. Inter-algal toxicity correlations

The relationship between the toxicity of 30 phenols to marine alga *D. tertiolecta* and freshwater alga *C. vulgaris* is provided in Figure 4.14.

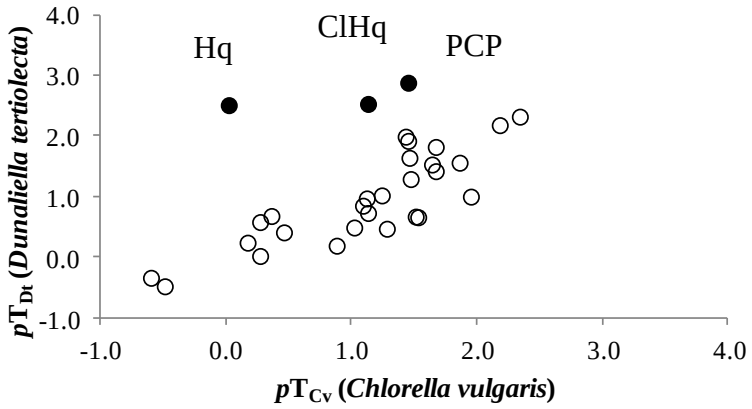


Figure 4.14. The plot of *C. vulgaris* versus *D. tertiolecta* toxicity data[3]

As illustrated in Figure 4.14, the inter-algal toxicity correlation suggests a definitive trend between two endpoints. However, as the visual analysis points out, the presence of statistical outliers, represented by black dots, decrease the relationship between the toxicity of phenols to marine and freshwater algae. When all the 30 phenols were analyzed, the linear regression analysis revealed a rather poor relationship between two algae (Eq. 4.8):

$$p\mathrm{T}\,(\mathrm{Dt}) = 0.72\,p\mathrm{T}\,(\mathrm{Cv}) + 0.34 \qquad (4.8)$$

(n=30, r^2=0.37, r^2_{cv}=0.27, *RMSE*=0.69)

As mentioned above, three statistical outliers were detected in the linear analysis. These compounds were hydroquinone, chlorohydroquinone and pentachlorophenol. The marine algal toxicity of hydroquinones, with the exception of tetrachlorohydroquinone, was severely underestimated by the inter-algal toxicity model where the freshwater toxicity data was used as the independent variable. It was interesting to note that the toxicity of TetraClHq was in good agreement between two algae. This is likely to be, as discussed earlier, due to the dominant polar narcotic mechanisms involved in the toxicity of

[3] 96 h toxicity for *C. vulgaris* and 48 h toxicity for *D. tertiolecta* based upon SGR and ICp calculations

TetraClHq towards algae. It was slightly more toxic to marine algae probably as a result of the absence of a rigid cell wall in *D. tertiolecta*, which in turn might have facilitated the transport of this chemical into the cytoplasm. Apart from TetraClHq, the other hydroquinones were much more toxic to marine alga as compared to freshwater alga. As discussed previously, the pronounced toxicity of hydroquinones is probably due to their transformation to more reactive quinone forms in the marine algal test system. The other compound that was an outlier to the relationship was pentachlorophenol (PCP), which was the most toxic phenol to marine alga, being classified as highly toxic according to European Union classification system (Table 4.2). However, it was classified as toxic to *C. vulgaris* (Table 4.10). The reason why PCP was more toxic to *D. tertiolecta* as compared to *C. vulgaris* is not clear. One possible explanation of the difference between the algal responses might be sought in the contribution of the ionized form of the molecule to toxicity. It seems likely that the ionized PCP in the marine algal test system is much more toxic to algal cells probably as a result of the absence of a cell wall in *D. tertiolecta*, which facilitates the transport of the molecule to the cytoplasm.

When the three compounds were removed from analysis, the inter-algal correlation was significantly improved and yielded much better results (Eq. 4.9):

$$pT\,(Dt) = 0.85\,pT\,(Cv) \qquad (4.9)$$

(n=27, r^2=0.72, r^2_{cv}=0.68, *RMSE*=0.39)

Note that the toxicity data set includes polar narcotics, respiratory uncouplers and several reactive chemicals even when the three outliers were removed. Yet, the intercept of Eq. 4.9 passes from zero, which suggests that the initial sensitivity of both test systems do not differ significantly. As suggested by Faber and Kowalski (1997), the *RMSE* values can be treated approximately as standard deviations of model predictions; as such, the *RMSE* summarizes the overall error of the model. Comparison of Eq. 4.8 and Eq. 4.9 suggest that the removal of outliers approximately halved the overall error of the inter-algal toxicity model between marine and freshwater alga. The error associated with Eq. 4.9 has *RMSE* (in log units) of 0.39. Furusjö et al. (2006) suggested that *RMSE* < 0.50 can be used as reliable indicators of a prediction for risk assessment purposes. However, the presence of

outliers, especially the different behavior of hydroquinones in marine and freshwater algal test systems requires attention when such a model is used for risk assessment purposes.

The use of freshwater data in marine hazard assessments is described in the European Union Technical Guidance Document (EU TGD, 2003), Canadian Water Quality Guidelines (CCME, 2007) and the USA marine water quality criteria (Russo, 2002). In using this surrogate approach the EU TGD (2003) calls for a clear understanding of the comparability of effects data generated on both types of species. The Canadian Water Quality Guidelines also allow the use of freshwater ecotoxicity data in marine risk assessments for the protection of aquatic life on a case-by-case basis in order to broaden the marine database (CCME, 2007). The United States Marine Quality Criteria also allow for freshwater data to be mixed with marine values particularly at the chronic level. However, there is no process for the surrogate approach (freshwater data *in lieu* of marine data) at either the acute or chronic levels (Russo, 2002).

As mentioned earlier, the EU TGD (2003) and CCME (2007) require argumentation on a case-by-case basis for the inclusion of freshwater data in marine hazard assessment. As stated by Bajot et al. (2011), the presence of a structural alert (i.e., presence of a di-hydroxy structure in the *para* configuration) in a chemical should be considered as a piece of information which should assist in the identification of the mechanism of chemical actions and hence grouping of chemicals. This kind of information can also be used to highlight compounds which should not be used in extrapolation from freshwater to marine effects. When the results from marine and freshwater toxicity data are evaluated, it can be concluded that the use of freshwater toxicity data of hydroquinones to *C. vulgaris* is not recommended to predict their toxicity to marine alga *D. tertiolecta.*

The literature reports regarding the toxicity of hydroquinone to several organisms suggest that the response of organisms to this chemical varies considerably. As can be seen from Table 4.16, the algae show a mixed response to hydroquinone: *D. tertiolecta, P. micans* and *P. subcapitata* are very sensitive to hydroquinone whereas *C. vulgaris* and *Crypthecodinium cohnii* are relatively more tolerant to this compound.

Table 4.16. Toxicity of hydroquinone to various aquatic organisms

Aquatic organism		EC_{50} (mg/L)	Reference
Algae	*Dunaliella tertiolecta*	0.33	This study
	Chlorella vulgaris	105.9	This study
	Pseudokirchneriella subcapitata	0.34[a]	Devillers et al. (1990)
	Crypthecodinium cohnii	50.0[a]	Devillers et al. (1990)
	Prorocentrum micans	0.30[a]	Devillers et al. (1990)
Protozoa	*Tetrahymena pyriformis*	37.0[a]	Cronin et al. (2002)
	Colpidium campylum	73.3[a]	Devillers et al. (1990)
Bacteria	*Vibrio fischeri*	0.038[a]	Devillers et al. (1990)
	Escherichia coli	34.0[a]	Devillers et al. (1990)
Macrophyta	*Elodea canadensis*	43.0[a]	Stom and Roth (1981)
	Lemna minor	7.7[a]	Stom and Roth (1981)
Crustacea	*Daphnia magna*	0.13[a]	Crisinel et al. (1994)

[a]: Transformed from *p*T values

The toxicity data on chlorohydroquinone on the other hand is relatively limited. Based on this limited information, this chemical was found to be particularly toxic to marine alga *D. tertiolecta* (Table 4.17). Similar to the case of hydroquinone, the sensitivity of marine algae to chlorohydroquinone as compared to other organisms (Table 4.17) suggest that extrapolation from freshwater to marine data should be carried out with large safety factors.

Table 4.17. Toxicity of chlorohydroquinone to various aquatic organisms

Aquatic organism		EC_{50} (mg/L)	Reference
Algae	*Dunaliella tertiolecta*	0.37	This study
	Chlorella vulgaris	10.7	This study
Protozoa	*Tetrahymena pyriformis*	8.0	Cronin et al. (2002)
Bacteria	*Vibrio fischeri*	5.8	TerraTox™ (2006)

The toxicity of pentachlorophenol is abundant in the literature and in general it was found to be very toxic to aquatic organisms (Table 4.18). Interestingly, the members of the genus *Chlorella* seems to be relatively tolerant to PCP compared to other algae.

Table 4.18. Toxicity of pentachlorophenol to various aquatic organisms

Aquatic organism		EC50 (mg/L)	Reference
Algae	*Dunaliella tertiolecta*	0.36	This study
		0.17	USEPA, 2012
	Chlorella vulgaris	9.4	This study
		10.3	Shigeoka et al. (1988)
		7.8	Repetto et al. (2001)
	Chlorella ovalis	5.5	Adema and Vink (1981)
	Pseudokirchneriella subcapitata	0.42	Shigeoka et al. (1988)
	Scenedesmus subspicatus	0.09	Geyer et al. (1985)
		0.18	Schäfer et al. (1994)
	Chlamydomonas reinhardi	0.22	Schäfer et al. (1994)
	Phaeodactylum tricornutum	3.0	Adema and Vink (1981)
Protozoa	*Tetrahymena pyriformis*	2.4	Cronin et al. (2002)
Bacteria	*Vibrio fischeri*	0.84	Crisinel et al. (1994)
	Escherichia coli	32.0	Nendza and Seydel, 1988
Crustacea	*Daphnia magna*	0.54	Repetto et al. (2001)

Based on the analysis of toxicity data on hydroquinone, chlorohydroquinone and pentachlorophenol, it can be concluded that toxicity of these chemicals to aquatic organisms varies considerably and extrapolation for these chemicals should be carried out with additional safety factors to be protective of marine algae.

4.3.2. MOA based evaluation

As suggested in the ECETOC workshop on the probabilistic approaches for marine risk assessment (2008), the mode of action (MOA) approach can be useful to analyze aquatic toxicity data and extract useful information for decision making. Therefore, the toxicity data sets were analyzed based on the MOA of phenols to investigate the confidence level that can be assigned to predictions when freshwater data is used to predict marine algal toxicity.

As illustrated in Figure 4.14, the visual inspection of the relationship between the toxicity of phenols to *D. tertiolecta* and *C. vulgaris* points towards some obvious patterns. It is of our interest to investigate whether it is possible to uncover these patterns underlying both toxicity data using the mechanism of toxic action approach. To this end, the inter-algal correlations were investigated based on the MOA of each chemical.

4.3.2.1. Polar narcotics. Partitioning of a hydrophobic compound into membranes causes disturbances in their structure and functioning, which, in turn, leads to the so-called baseline toxicity, or narcosis. This constitutes the minimal toxicity of any hydrophobic pollutant (Escher and Schwarzenbach, 2002). Narcosis (i.e., non-polar and polar narcosis) is the least specific, but arguably the most important mode of toxic action in ecotoxicology because approximately 70% of all organic industrial chemicals are believed to act via narcosis (Bradbury and Lipnick, 1990). Therefore, an inter-algal toxicity correlation for polar narcotic compounds is of utmost importance considering the number of chemicals categorized under this MOA.

In their critical review on marine risk assessment, Peters et al. (2005) conducted an analysis of a large database covering compounds with narcotic (MOA1) and polar narcotic (MOA2) mode of actions according to the classification by Verhaar et al. (1992) to investigate extrapolation viability between marine and freshwater species. For 20 narcotic compounds, the researcher concluded that only 20% of the variation in marine algal toxicity was explained by toxicity data on freshwater algae. As for polar narcotics, the researchers reported that the number of marine algal toxicity data were limited and the available data on the freshwater and saltwater toxicity of algae (n=8) showed no functional relationships. However, as also stated by the researchers, their study on the relative sensitivity of freshwater and marine organisms suffered from a high level of background-variability because the data was not produced for such a comparison. In contrast, the data generated in this study provided the unique opportunity to compare the responses of marine and freshwater algae, because the data was generated for such a comparison; therefore should provide a realistic assessment with minimum background variability.

The visual inspection of the relationship revealed that the toxicity of 19 polar narcotics was in good agreement between two algal species (Figure 4.15). Accordingly the linear regression analysis resulted in a strong trend:

$$pT\,(Dt) = 0.83\,pT\,(Cv) \qquad (4.10)$$

(n=19, r^2=0.82, r^2_{cv}=0.78, $RMSE$=0.30)

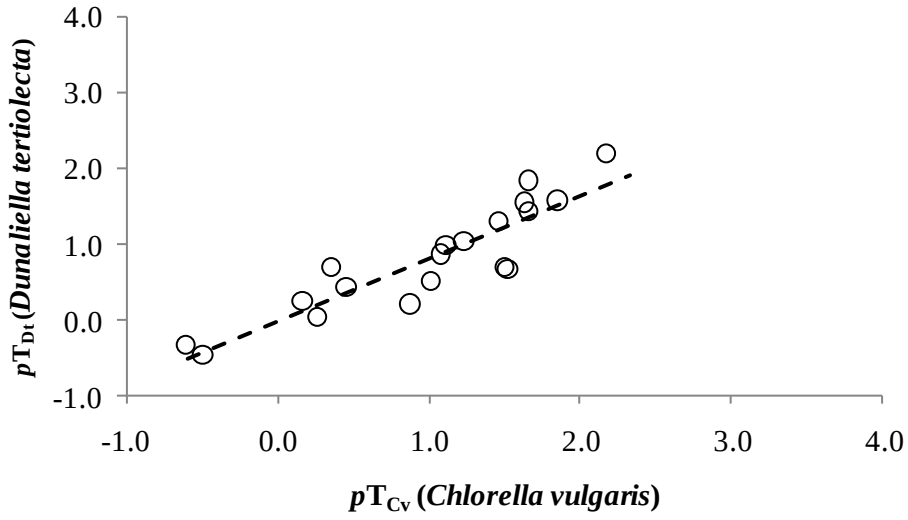


Figure 4.15. The relationship between the algal toxicity of polar narcotics to *C. vulgaris* and *D. tertiolecta*

The relatively low error associated with Eq. 4.10 (*RMSE*=0.30) suggests that marine algal toxicity of polar narcotics can be predicted with confidence using toxicity data on freshwater algae. Further investigation of the relationship revealed that 2,6-DCP, 2,3,6-TCP and 2,4,6-TCP were less toxic to marine alga as compared to freshwater alga. As discussed previously, the simultaneous occupation of both *ortho* positions renders these chlorophenols less toxic compared to other congeners with similar number of chlorine substitution. It seems likely that the chlorophenols with both *ortho* positions occupied are markedly less toxic in the marine algal test system as compared to freshwater system. This might be due to the presence of excess chlorine atoms in the marine test medium which might have led to an increased hindrance, hence reduced toxicity, in the chlorophenol molecule. Another reason might be the inter-species differences. No matter what the reason, the use of freshwater data on *C. vulgaris* would over-estimate their toxicity to marine alga *D. tertiolecta*, which is preferable to under-estimation with regard to risk assessment (Cronin et al., 2002). Removal of these three compounds produced a much better correlation and lower error (Eq. 4.11):

$$p\text{T (Dt)} = 0.89\ p\text{T (Cv)} \qquad (4.11)$$

(n=16, r^2=0.93, r^2_{cv}=0.91, *RMSE*=0.20)

The error of the relationship presented in Eq. 4.11 was reduced by half, as compared to the one presented in Eq. 4.10. Since the outliers are easily rationalized from a chemical perspective, this relationship can be used to extrapolate toxicity of polar narcotics between *D. tertiolecta* and *C. vulgaris*.

In an attempt to develop a model that predicts with confidence the toxicity of polar narcotics without an outlier, several molecular descriptors were included in the analysis to improve the predictivity of the relationship reported in Eq. 4.10. The following model was found to be highly significant, predictive and interpretable from a toxicological perspective (Eq. 4.12):

$$p\text{T (Dt)} = 1.10\ (\pm 0.07)\ p\text{T (Cv)} + 0.30\ p\text{Ka}\ (\pm 0.05) - 2.65\ (\pm 0.47) \quad (4.12)$$

(n=19, r^2=0.94, r^2_{cv}=0.92, *RMSE*=0.17)

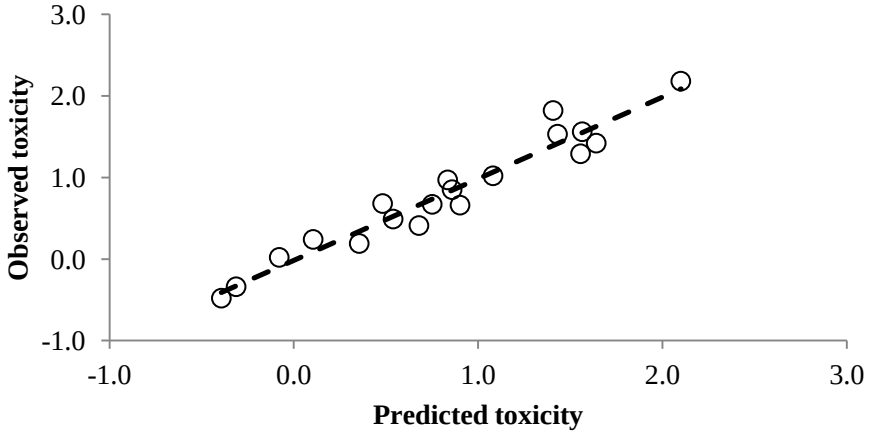


Figure 4.16. Plot of observed toxicity versus toxicity predicted using Eq. 4.12.

The observed and predicted toxicity was in excellent agreement after the inclusion of *p*Ka into the equation (Figure 4.16). The *p*Ka parameter carries significant physical information regarding the tendency of phenols to ionize at a particular *p*H. The Eq. 4.12 concurs with our previous findings that the ionization of phenols in the freshwater algal test system employing *C. vulgaris*, described by the right-hand side of the equation, is a crucial process that should be taken into consideration when toxicity data is to be used for extrapolation. Accordingly, correcting for the term describing the deprotonation tendency of the molecules, toxicity data generated in both algal test systems were found to be in

excellent agreement with each other which also resulted in very low model error (*RMSE*=0.17).

4.3.2.2. Respiratory uncouplers. Weak acid respiratory uncouplers disrupt ATP synthesis and are generally bulky and electronegative. They cause the inner mitochondrial membrane to be permeable to hydrogen ions, thereby disrupting the hydrogen ion gradient (Schultz, 1999). From a predictive perspective, the aquatic toxicity of uncouplers was usually found to be in good agreement with their hydrophobicity, described either by Log *P* or Log *D*. Using several QSTR models employing a hydrophobicity term and an electronic parameter, Fujita (1966) concluded that only the neutral species is physiologically active. In contrast, Tabata (1962) postulated that the neutral phenol and the deprotonated phenoxide are both active and exhibit additive effects but he argued that the neutral species is more toxic than the charged species, mainly because of its higher uptake. Könemann and Musch (1981) developed a QSTR equation with Log *P* and *p*Ka as descriptors, while emphasizing that such a QSTR is hampered by the co-linearity of these descriptors for the test set of chlorophenols. The QSTR analysis for the algal toxicity data sets generated in this study revealed that while marine algal toxicity of the respiratory uncouplers were in good agreement with Log *P*, the toxicity of the same chemicals to freshwater alga *C. vulgaris* had a good correlation only to the *p*H corrected hydrophobicity, Log *D*. As for the interspecies toxicity correlation, the analysis revealed that there was almost no relationship between the responses of marine and freshwater algae to respiratory uncouplers (*n*=4, r^2=0.05).

As discussed previously, the toxicity of respiratory uncouplers to freshwater alga *C. vulgaris* depends highly on the test *p*H. Weak acids, such as chlorophenols, tend to ionize at a *p*H greater than their acid dissociation constants (*pK*a) and the degree of ionization enhances as the (*p*H – *pK*a) difference increases. The decrease in toxicity of weak acids has been attributed to the fact that unionized form of the molecule contributes to the toxicity more than the ionized form because the neutral molecule is more bioavailable than the corresponding charged molecule (Escher and Schwarzenbach, 2002). Note that the *p*Ka of 2,3,5,6-TeCP, 2,3,4,6-TeCP and PCP are lower than the test *p*H (Table 4.10). It is likely that ionization rendered pentachlorophenol and these two tetrachlorophenols (i.e., 2,3,5,6-TeCP and 2,3,4,6-TeCP) less toxic. This also explains why 2,3,4,5-tetrachlorophenol was

more toxic considering that the *p*Ka of this compound is above the test *p*H (Table 4.11), hence it should be more bioavailable to algae during the bioassays compared to other TeCPs and PCP. Another reason of higher toxicity of this compound may result from the unoccupied *ortho* position (i.e., 6- position) in 2,3,4,5-TeCP as opposed to two other TeCPs and PCP, which have both *ortho* positions occupied by chlorine atoms.

Based on the explanations above, the reason behind the low correlation among the toxicity data for respiratory uncouplers to *D. tertiolecta* and *C. vulgaris* seems to be closely related to the transport of these chemicals to the inner membranes where they disrupt the ATP synthesis. Due to the low number of compounds, a QSTR analysis was not attempted. However, it should be underlined that the different behavior of uncouplers in marine and freshwater algal test systems was best captured by the hydrophobicity parameters, Log *P* (Eq. 4.1) and Log *D* (Eq. 4.4), which highlights the relative role of ionization on these two algal systems.

4.3.2.3. Reactive (electrophilic) phenols. Electrophiles, often referred to as reactive toxicants, are associated with irreversible interaction with biological macromolecules. Some phenols may be directly electrophilic, and others may become active through tautomeric keto forms. In some cases, the toxicity is driven from the chemical properties of other substituents rather than from the phenolic group (Aptula et al, 2005b). From a chemical perspective, electrophiles *in vivo* target nucleophilic sites in peptides, proteins, and nucleic acids. Interactions between reactive chemicals and cellular components are not specific, but they can disrupt many different cellular processes. These so-called ‘‘reactive chemicals’’ are responsible for a wide variety of adverse outcomes. Modeling these adverse outcomes is a serious challenge because many chemicals are ‘‘activated’’ by metabolic processes (Schultz et al., 2006). As such, the difficulties in modeling the aquatic toxicity of compounds that may act either directly or indirectly as electrophiles have been reported in various studies (Cronin et al., 2002; Enoch et al., 2008; Ertürk et al., 2012).

The inter-algal toxicity correlation for the reactive phenols tested in this thesis turned out to be rather poor (n=7, r^2=0.05) due mainly to the distinct behavior of hydroquinones in two different algal test systems. In an attempt to improve the relationship between the

two toxicity data, several molecular descriptors were involved in the relationship. The following relationships were found to be highly significant (Table 4.19):

Table 4.19. Reactivity models developed for *D. tertiolecta* using *C. vulgaris* toxicity data and other molecular descriptors

Model ID	Model	(*n*=7)			Descriptor information
		r^2	r^2_{CV}	*RMSE*	
R1	pTDt=0.64 $p\mathrm{T_{Cv}}$ - 18.18 EEig07r - 9.14	0.98	0.94	0.23	Eigenvalue descriptor
R2	pTDt=0.60 $p\mathrm{T_{Cv}}$ - 6.90 BEHe7+ 8.94	0.98	0.95	0.19	Burden eigenvalue
R3	pTDt=0.43 $p\mathrm{T_{Cv}}$ + 335.20 $\mathrm{NRI_{minC}}$	0.92	0.85	0.31	Minimum nucleophilic reactivity index for atom Carbon

The models describing the relationship between marine and freshwater algae for reactive phenols are complemented with an electronic descriptor in each model. In the first model, the difference in the reactivity of phenols, in particular hydroquinone and chlorohydroquinone, is described by EEig07r, which is an edge adjacency index weighted by resonance integrals. In the second model, BEHe7, which is a burden eigenvalues descriptor obtained by weighing atomic Sanderson electronegativity of the atoms, covers the differences in the responses by two algae. Both EEig07r and BEHe7 were obtained from Dragon 5.4 and by definition they are related to the electronegativity of the compound. In the third model, the nucleophilic reactivity index for carbon atom was obtained from CODESSA Pro and is defined as $\sum_{i\varepsilon A} C^2_{iHOMO}$, where the summation is performed over all atomic orbitals *i* of a given atom *A*, and $C_{i\mathrm{HOMO}}$ denotes the i^{th} atomic orbital coefficient for the HOMO. These indexes give an estimate of the relative reactivity of the atoms in the molecule and are related to the activation energy of the corresponding chemical reaction (Jover et al., 2009). Thus, the minimum nucleophilic reactivity index for carbon atom is believed to describe the bonding ability of the phenol to possible nucleophiles.

The models clearly indicate that the reactivity of phenols towards marine and freshwater algae required the inclusion of an electronic descriptor to account for the different responses. The reason behind the difference in the algal responses might be due to the biological differences of algae as well as differences in the test medium as it has been discussed before.

4.3.3. Global models

To develop global QSTRs that can predict the toxicity of compounds without an outlier to marine alga *D. tertiolecta* using toxicity data on *C. vulgaris* and molecular descriptors, several models were constructed using the training set compounds and their predictivity was assessed based on the predictions performed on the test set compounds. For this purpose, the training/test set division that was employed for *D. tertiolecta* global QSTR models as outlined in section 4.1.2.1 was used. This division enabled the comparison of models employing toxicity data on *C. vulgaris* as the forced-in descriptor to those developed previously in section 4.1.2.1. Results revealed that the attempts to develop a better model failed as none of the models were statistically superior over the ones developed previously for *D. tertiolecta*. This was in great part due to the different response of marine and freshwater algae to Hq, ClHq and PCP which resulted in the insignificance of *C. vulgaris* data as the independent variable. Therefore, the QSTR analysis was conducted following the removal of Hq, ClHq and PCP. In the reduced data sets, the relationships presented in Eq. 4.13 resulted in a quality QSTR (which was developed using the training set in Table 4.3 without PCP and ClHq and its predictivity was assessed using the compounds in the test set except hydroquinone).

$$pT\,(Dt) = 0.32\,pT\,(Cv) + 13.48\,Mor30p - 1.31 \qquad (4.13)$$

$n_{tr}=22$, $r^2=0.88$, $r^2_{cv}=0.85$, $RMSE_{tr}=0.26$

$n_{test}=5$, $Q^2_{F3}=0.76$, $R^2=0.78$, $RMSE_{te}=0.35$

As can be seen from the statistical summary of the inter-algal QSTR presented in Eq. 4.13, the use of freshwater toxicity data together with a 3D-MoRSE descriptor, namely, Mor30p (3D-MoRSE – signal 30 / weighted by atomic polarizabilities), resulted in a highly significant and predictive QSAR. It should be noted that the model also met the external predictivity criteria by Golbraikh et al. (2003) as outlined in Eqs. 3.4-3.7 under Materials and methods ($R_0^2=0.71$, $k=1.10$, $R_0'^2=0.82$, $k'=0.78$).

Although the statistical summary provided together with Eq. 4.13 is in agreement with the principles of QSTR construction (i.e., construction of a model using training set

compounds and evaluation of its predictive power by using a test set), it is also appropriate to provide a statistical analysis of the whole data set. Using the same two descriptors for all compounds except Hq, ClHq and PCP, the following relationship was obtained:

$$pT\,(Dt) = 0.31\,pT\,(Cv) + 13.05\,Mor30p - 1.21 \quad (4.14)$$

(n=27, r^2=0.86, r^2_{cv}=0.82, *RMSE*=0.27)

As suggested by Eq. 4.14, the toxicity data on *C. vulgaris* complemented with a 3D-MorSE descriptor was able to explain the 86% variance in the toxicity of 27 phenols to *D. tertiolecta*. It should be stated that the aim of the QSTR analysis using *C. vulgaris* toxicity data as the forced-in independent variable in the modeling process was not to develop a predictive model but rather to use QSTR methodology to carry out a comparative analysis on algal toxicity data generated in the same laboratory, by the same researcher and according to standardized test protocols. From this perspective, the relationship provided above indicated that to predict the toxicity of phenols to *D. tertiolecta* using toxicity data on freshwater alga *C. vulgaris*, the model required an additional 3D-MoRSE descriptor, namely, Mor30p. As discussed previously, 3D-MoRSE descriptors provide information of a molecule on its 3D structure by taking several atomic properties into account (Mercader and Pomilio, 2010). Mor30p, in this respect, can be said to correct for the polarizability difference of phenols in marine and freshwater test systems. The polarizability difference might occur because the rates of many organic reactions, including those between electrophiles and nucleophiles, are significantly influenced by the polarity of the reaction medium (Aptula and Roberts, 2006).

4.3.4. Interspecies toxicity correlations using literature toxicity data

The measured toxicity for marine alga *D. tertiolecta* and *C. vulgaris* were correlated to the acute toxicity to aquatic species such as the freshwater alga *Pseudokirchneriella subcapitata* (formerly known as *Selenastrum capricornutum*), marine bacterium *Vibrio fischeri (*formerly known as *Photobacterium phosphoreum)*, freshwater protozoan *Tetrahymena pyriformis*, freshwater daphnid *Daphnia magna*, and freshwater fish *Pimephales promelas* to yield interspecies toxicity correlations. Although there are some

gaps in the toxicity values, a relatively large number of data for comparison were collected for the majority of organisms (Table 4.20).

Table 4.20. The toxicity data used to construct interspecies correlations

Compound	*p*T (Dt) (48 h)	*p*T (Cv) (96 h)	*p*T (Ps) (72 h)	*p*T (Vf) (30 min)	*p*T (Tp) (40 h)	*p*T (Dm) (48 h)	*p*T (Pp) (96 h)
P	-0.34	-0.60	-0.32	0.42	-0.21	0.51	3.46
2-CP	0.24	0.17	0.39	0.58	0.18	0.86	4.05
3-CP	0.68	0.36	1.05	0.96	0.87	0.91	n.a.
4-CP	0.41	0.46	0.61	1.19	0.55	1.20	4.32
2,3-DCP	0.85	1.09	1.17	1.52	1.28	1.50	n.a.
2,4-DCP	1.02	1.24	1.30	1.47	1.04	1.78	4.32
2,5-DCP	0.97	1.12	1.65	1.24	1.13	1.56	n.a.
2,6-DCP	0.19	0.88	1.01	1.09	0.73	1.24	n.a.
3,4-DCP	1.29	1.47	1.87	2.00	1.75	1.77	n.a.
3,5-DCP	1.42	1.67	1.89	1.62	1.57	1.89	n.a.
2,3,4-TCP	1.53	1.64	1.68	2.20	n.a.	1.95	n.a.
2,3,5-TCP	1.56	1.86	1.94	2.25	2.37	1.94	n.a.
2,3,6-TCP	0.67	1.51	1.39	1.19	n.a.	1.43	n.a.
2,4,5-TCP	1.82	1.67	1.42	2.19	2.10	1.98	n.a.
2,4,6-TCP	0.66	1.53	1.54	1.41	1.41	1.56	4.61
3,4,5-TCP	2.18	2.18	n.a.	2.74	n.a.	2.35	n.a.
2,3,4,5-TCP	2.32	2.34	n.a.	3.11	2.71	2.12	5.75
2,3,4,6-TCP	1.92	1.45	2.25	2.26	2.18	1.93	5.35
2,3,5,6-TCP	1.99	1.43	n.a.	2.02	2.22	2.01	n.a.
PCP	2.88	1.45	2.80	2.71	2.05	2.54	6.08
Pyrogallol	0.47	1.28	n.a.	n.a.	0.85	n.a.	n.a.
Hq	2.51	0.02	2.51	3.46	0.47	2.93	n.a.
ClHq	2.53	1.13	n.a.	1.40	1.26	n.a.	n.a.
TetraClHq	1.64	1.46	n.a.	n.a.	2.11	n.a.	n.a.
Cat	0.58	0.27	n.a.	0.57	0.75	n.a.	4.08
4ClCat	0.73	1.13	n.a.	n.a.	1.06	n.a.	n.a.
35DiClCat	1.00	1.95	n.a.	2.08	1.37	n.a.	n.a.
Res	-0.48	-0.49	n.a.	-0.38	-0.65	0.30	n.a.
4ClRes	0.02	0.27	n.a.	n.a.	0.13	n.a.	n.a.
46DiClres	0.49	1.02	n.a.	n.a.	0.97	n.a.	n.a.

n.a., data not available

4.3.4.1. Correlations between *D. tertiolecta* and aquatic organisms. The measured toxicity values for *D. tertiolecta* were correlated with the acute toxicity to alga *P. subcapitata*, bacterium *V. fischeri*, protozoan *T. pyriformis*, daphnid *D. magna* and fish *P. promelas*. Although there were some gaps in the toxicity values, a relatively large number of data for comparison were collected. A plot of the correlations between toxicity to the five species is presented in Figure 4.17.

Figure 4.17 revealed that the marine algal toxicity of phenols to *D. tertiolecta* correlated strongly with toxicity data on other aquatic species. The relationships between *D. tertiolecta* and freshwater alga, freshwater daphnid and freshwater fish were exceptionally good and did not produce an outlier. The correlations between the toxicity of phenols to marine alga and marine bacterium and freshwater protozoan were also highly satisfactory. However, some outliers were detected to these relationships. It was found that hydroquinone, chlorohydroquinone and pentachlorophenol elicited different responses by these organisms (Table 4.21).

Table 4.21. Interspecies toxicity correlations between *D. tertiolecta* and five aquatic species

Species	**Model**[a]	***n***	r^2	r^2_{cv}	***RMSE***	**Outlier(s)**
P. subcapitata	1.00 *p*T (Ps) – 0.33	18	0.84	0.79	0.32	-
V. fischeri	0.90 *p*T (Vf) – 0.34	24	0.87	0.85	0.31	ClHq
T. pyriformis	0.84 *p*T (Tp) – 0.10	24	0.93	0.92	0.18	Hq, ClHq, PCP
D. magna	1.36 *p*T (Dm) – 1.04	22	0.90	0.87	0.28	-
P. promelas	1.20 *p*T (Pp) – 4.52	9	0.96	0.94	0.21	-

[a] The models were constructed following the removal of outliers, if any.

The slopes of the relationships between *D. tertiolecta* and other aquatic organisms were close to unity, except that of *D. magna* and to some extent, that of *P. promelas* (Table 4.21). This finding indicates a one-to-one toxicity relationship of marine alga to marine bacterium, freshwater protozoan and freshwater algae, probably as a result of similar toxic modes of action of phenols towards these organisms. The large intercepts (>1) are probably due to the differences in the complexity of test organisms (unicellular vs. complex) which was clearly reflected in the interspecies toxicity relationships for freshwater daphnid and fish.

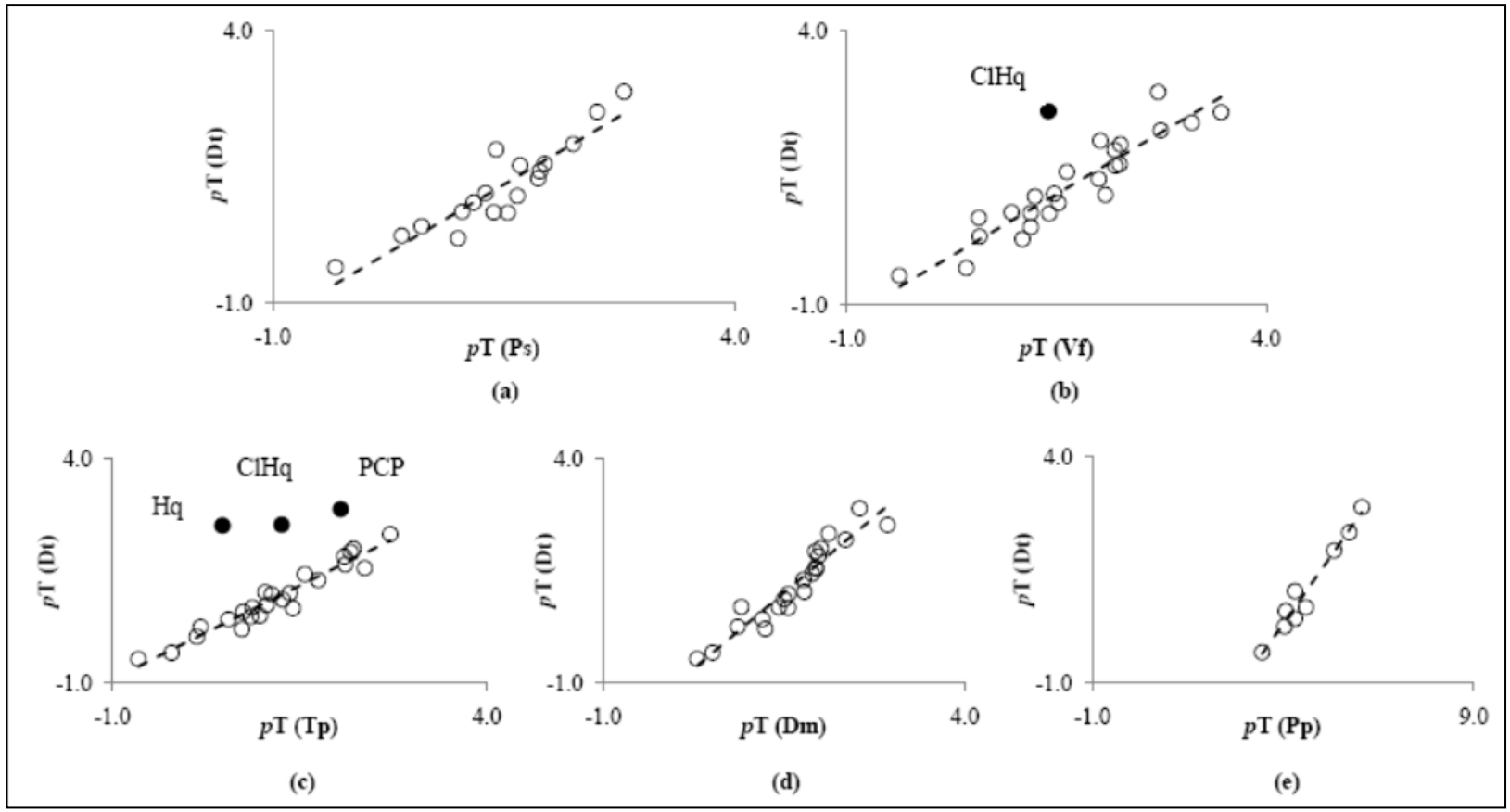


Figure 4.17. Interspecies toxicity relationships constructed between *D. tertiolecta* and a) *P. subcapitata*, b) *V. fischeri*, c) *T. pyriformis*, d) *D. magna* and e) *P. promelas* (black dots indicate statistical outliers)

Closer examination of the interspecies toxicity correlation between two algal species revealed a highly satisfactory relationship between *D. tertiolecta* and *P. subcapitata* (Table 4.21). The exact unity of the slope suggests that the two organisms respond similarly to phenolic compounds. The good correlation of the toxicities of these compounds to *D. tertiolecta* and *P. subcapitata* suggest that it is possible to use toxicity data on freshwater algae *in lieu* of marine algal toxicity.

As for the relationship between toxicity of phenols to marine alga and other aquatic species, the strong interspecies toxicity correlation observed between *D. tertiolecta* and *V. fischeri* suggests that the available toxicity data on *V. fischeri* can be used with confidence within the applicability domain of the respective model to predict marine algal toxicity of chemicals (Table 4.21). However, the different behavior of ClHq in the two test system reduces the reliability of this relationship for compounds expected to act through reactive mechanisms should it be used in a risk assessment perspective without MOA classification. Considering that there is extensive toxicity data for *V. fischeri* are available in the form of monographs (Kaiser and Palabrica, 1991) or in commercial databases (TerraTox™, 2006) covering over 1200 chemicals, the interspecies toxicity model developed in this thesis has the potential to predict the toxicity of chemicals, in particular polar narcotics and respiratory uncouplers, available to *V. fischeri* and not to *D. tertiolecta*.

The relationship between *D. tertiolecta* and *T. pyriformis* was found to be highly satisfactory only following the removal of three outliers from analysis. Note that the compounds that appeared as outliers to this relationship were also found to elicit different responses by *C. vulgaris* and *D. tertiolecta*. *Daphnia magna* is one of the species to which toxicity data must be provided within REACH. Therefore, the interspecies toxicity relationships between algae and daphnia can be used in efforts to extrapolate between the toxicity data on these organisms. Based on the relationship between *D. magna* and *D. tertiolecta*, the initial findings suggest that there is a good extrapolation possibility; however, the lack of reactive toxicity data on *D. magna*, except Hq, does not allow a thorough assessment across different MOAs. Nevertheless, for compounds expected to act through polar narcosis and respiratory uncoupling, preliminary findings point towards a reliable extrapolation. As for the relationship between toxicity of phenols to marine alga and freshwater fish, the intercept of the equation suggests that fish test system is much

more sensitive to phenols than marine alga *D. tertiolecta*. Although the relationship is excellent statistically; the low number of data hampers a definitive conclusion as to reliability of a possible extrapolation between the toxicity data on these two organisms.

4.3.4.2. Correlations between *C. vulgaris* and aquatic organisms. To investigate the interspecies relationships, the measured toxicity values for *C. vulgaris* were correlated with the acute toxicity to five aquatic species. A plot of the correlations between toxicity to the five species is presented in Figure 4.18.

Figure 4.18 revealed that the interspecies toxicity correlations between *C. vulgaris* and *P. subcapitata*, *V. fischeri* and *D. magna* were highly satisfactory following the removal of Hq which was detected as a common outlier in these relationships.

The slopes of the relationships between *C. vulgaris* and other aquatic organisms, except the one between *C. vulgaris* and *D. magna*, were close to unity, implying a one-to-one toxicity relationship between algae and other aquatic organisms. The large intercepts (>1) observed in the relationships between *C. vulgaris* and *D. magna* and *P. promelas*, as mentioned before, probably reflect the differences in the complexity of test organisms (unicellular vs. complex).

Table 4.22. The interspecies toxicity correlations between *C. vulgaris* and five aquatic species

Species	**Model**[a]	***n***	r^2	r^2_{cv}	***RMSE***	**Outlier(s)**
P. subcapitata	0.95 *p*T (Ps) – 0.14	16	0.83	0.78	0.40	Hq, PCP
V. fischeri	0.84 *p*T (Vf) – 0.17	24	0.81	0.77	0.33	Hq
T. pyriformis	0.80 *p*T (Tp)	27	0.81	0.78	0.33	-
D. magna	1.23 *p*T (Dm) – 0.80	21	0.84	0.78	0.31	Hq
P. promelas	0.88 *p*T (Pp) – 3.19	9	0.72	0.48	0.45	-

[a] The models were constructed following the removal of outliers, if any.

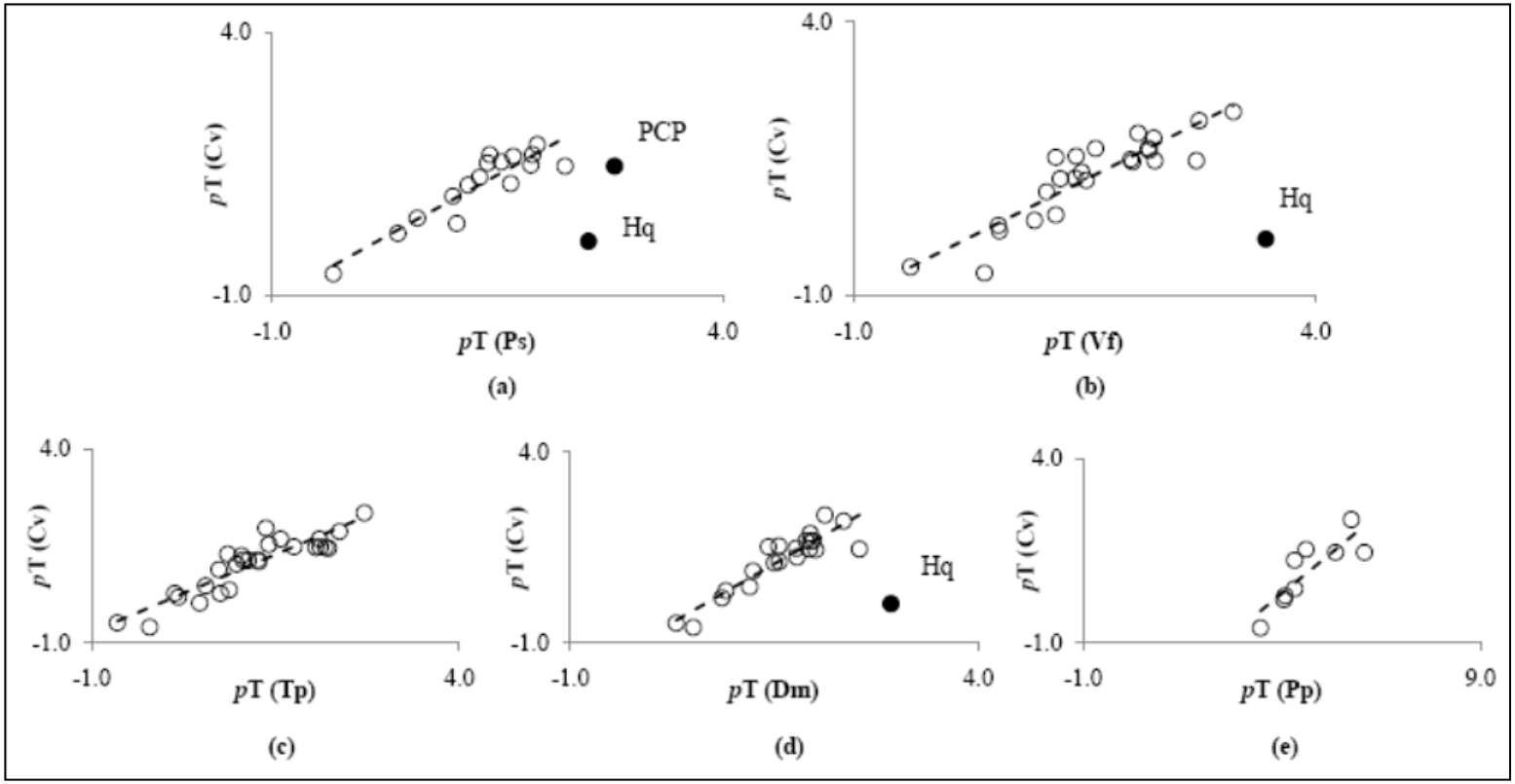


Figure 4.18. Interspecies toxicity relationships constructed between *C. vulgaris* and a) *P. subcapitata*, b) *V. fischeri*, c) *T. pyriformis*, d) *D. magna* and e) *P. promelas* (black dots indicate statistical outliers)

The relationship between the toxicity of phenols to *C. vulgaris* and freshwater protozoan was particularly promising considering the absence of an outlier and goodness of the fit (Table 4.22). *Tetrahymena pyriformis* toxicity data set is particularly important because of its historical use in QSTR studies and its high repeatability (Hewitt et al., 2011). The promising interspecies toxicity correlation suggest that historical *T. pyriformis* toxicity data set can be used in efforts to predict toxicity of untested phenols to freshwater alga *C. vulgaris*. The fact that there was no outlier in the relationship and relatively low error of the model (Table 4.22) suggests that this relationship can be exploited in attempts to predict the toxicity of a huge number of compounds (approximately 2400) to *C. vulgaris* with confidence for compounds within the applicability domain of the interspecies toxicity correlation developed in this thesis.

Previously, Cronin et al. (2004) developed a toxicity data set of 91 diverse chemicals towards *C. vulgaris* using the methodology developed by Worgan et al. (2003), which is a 15-minute assay, based on the premise that all organisms, including algae, contain non-specific esterases, the activity of which can be assessed by the measurement of the disappearance of an ester, or the appearance of the product. In their study, Cronin et al. (2004) demonstrated the viability of an interspecies toxicity correlation between these two organisms. However, the algal endpoint they used for the correlations is not compatible with the requirements of REACH for toxicity data on algae. In this thesis, it was demonstrated for the first time that such a relationship is also viable for the REACH compatible algal endpoint (i.e., growth inhibition).

The interspecies toxicity relationship between *C. vulgaris* and *P. subcapitata* merits further analysis. As can be seen from Table 4.22, the slope of the interspecies toxicity correlation was close to unity, indicating similar responses to phenols by *C. vulgaris* and *P. subcapitata*, but only following the removal of Hq and PCP from the relationship. Examination of the data set revealed that the remaining phenols were polar narcotics except 2,3,4,6-TeCP (Table 4.20). When the inter-algal correlations for the three algal species were considered, there seems to be a definitive trend for polar narcotic phenols among algae. Indeed, as the visual analysis suggests (Figure 4.19), the toxicity of polar narcotics to three

algae (*D. tertiolecta*, *C. vulgaris* and *P. subcapitata*) were highly correlated to hydrophobicity as described by Log *P*.

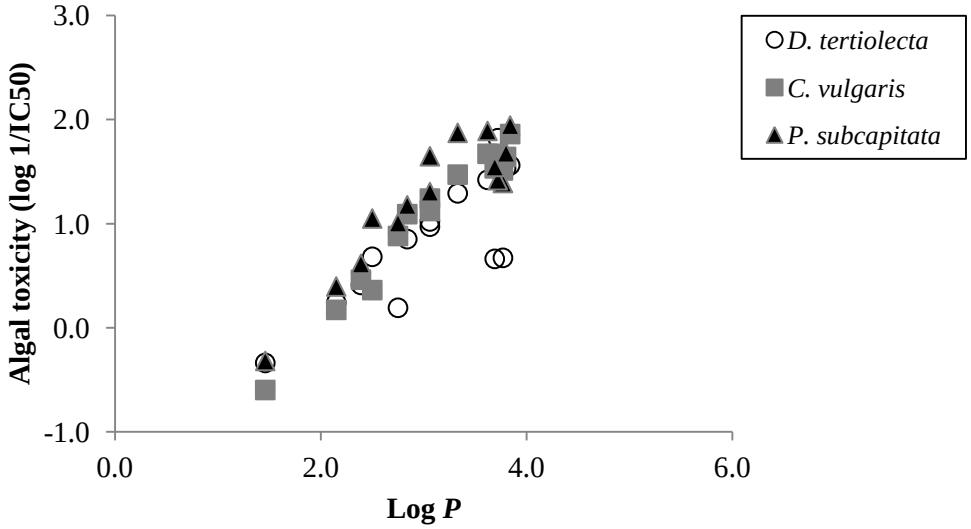


Figure 4.19. The relationship between polar narcotic toxicity to algae and hydrophobicity

The results imply that the toxicity data of polar narcotic compounds on algae can be used inter-changeably when data on one species is lacking. The excellent inter-algal correlations for polar narcotics but rather poor relationships for respiratory uncouplers and reactive compounds also suggest that chemical toxicity to algae differentiates above the level of simple membrane perturbations. In conclusion, the relationship between *D. tertiolecta* and *P. subcapitata* was found to be highly significant and did not result in any outliers; whereas the relationship between *C. vulgaris* and *P. subcapitata* resulted in two outliers, namely PCP and Hq. As discussed previously, the *C. vulgaris* test system has high *p*H dependence and the toxicity of PCP is relatively low due to issues related to ionization. Additionally, as can be seen from Table 4.18, the genus *Chlorella* seems to be relatively tolerant to PCP compared to other algae. As for hydroquinone, the response of aquatic organisms changes considerably to this chemical; therefore, its toxicity should ideally be determined in the laboratory.

As for the relationship between toxicity data on *C. vulgaris* and other aquatic organisms, the following conclusions can be drawn: similar to the relationship between *D. tertiolecta* and

V. fischeri, the interspecies toxicity correlation between *C. vulgaris* and marine bacterium has the potential to be used for filling the data gaps in algal toxicity, but the different behavior of hydroquinone in the respective test systems requires attention when algal toxicity of reactive compounds is to be predicted using data on *V. fischeri*. The interspecies correlation between *C. vulgaris* and *D. magna* was relatively unreliable due to the different responses of the organisms to hydroquinone. Nevertheless, following the removal of Hq, an acceptable relationship was obtained. Bearing in mind the possibility of different responses to reactive toxicants, in particular the hydroquinones, the interspecies toxicity relationship developed between *C. vulgaris* and *D. magna* in this study has the potential to be used for extrapolation purposes when data on one species is lacking. As for the relationship between *C. vulgaris* and *P. promelas*, the analysis revealed that the presence of PCP was the main reason of the poor relationship, although it was not detected as a statistical outlier. The removal of PCP from analysis significantly increased the interspecies correlation (n=8, r^2=0.85, r^2_{cv}= 0.77). As mentioned before, the uncouplers are strongly influenced by the *p*H of the freshwater algal test system. On the basis of this finding, an interspecies toxicity correlation between *C. vulgaris* and *P. promelas* is not recommended to be used in a risk assessment perspective, at least for a data set comprised of compounds acting through different MOAs. On the other hand, for polar narcotic compounds, the relationship between *C. vulgaris* and *P. promelas* was noted to be excellent and it should be safe to use toxicity data inter-changeably for these compounds.

4.4. Risk assessment perspective

In Europe, due to the new chemical policy, REACH, QSAR models have become important tools since thousands of chemicals have to be assessed (Comber et al., 2003). The most extensively validated and used QSAR is the ECOSAR which is an easy-to-use computer programme developed and routinely applied by the United States Environmental Protection Agency for predicting aquatic toxicity to fish, aquatic invertebrates (daphnids), and green algae (Reuschenbach et al., 2008).

In this part of the thesis, the widely used ECOSAR software was tested in its predictivity (section 4.4.1) and classification ability (section 4.4.2) using the experimental data set of

phenols generated in this study for marine and freshwater algae. The algal data sets generated in this thesis provided the unique opportunity to test the prediction capability of ECOSAR using data generated in a single laboratory by the same researcher according to standardized algal growth inhibition tests.

4.4.1. ECOSAR predictions

The ECOSAR prediction for the endpoint and toxicity class is provided in Table 4.23. For phenols there are different ECOSAR classes that can be used for predictions. Irrespective of these classes, the ECOSAR uses available toxicity data stored in the ECOSAR database for constructing the hydrophobicity-based model to be used in the predictions. It was found that for 8 phenols tested in this study (pyrogallol, chlorohydroquinone, tetrachlorohydroquinone, catechol, 4-chlorocatechol, 3,5-dichlorocatechol, 4-chlororesorcinol and 4,6-dichlororesorcinol), there was no toxicity data available for algae in the ECOSAR database. In this respect, the algal toxicity data generated in this study makes an important contribution by providing new toxicity data of 8 chemicals on two different algae which is not available in the ECOSAR database (USEPA, 2011).

Based on the results obtained from ECOSAR predictions, it was found that for polar narcotic chemicals, the use of neutral organic class produced the best results for both algae (Table 4.23). For respiratory uncouplers, the use of neutral organic class also produced highly satisfactory results for marine alga *D. tertiolecta*, while it over-estimated the toxicity of uncouplers to freshwater alga *C. vulgaris*. This is believed to be related to the ionization phenomenon that renders uncouplers less toxic in the *C. vulgaris* test system than predicted by the ECOSAR model. The lack of a cell wall in marine alga *D. tertiolecta* might also be another reason as to why the toxicity of uncouplers converges on a narcosis mechanism as evidenced by the excellent correlation of Log *P* with toxicity (Eq. 4.1). On the other hand, an equally good relationship was obtained only when the *p*H corrected hydrophobicity, Log *D*, was used to model the toxicity of uncouplers to *C. vulgaris* (Eq. 4.4).

Table 4.23. ECOSAR predictions for *D. tertiolecta* and *C. vulgaris*

Compound	Expected MOA	*p*T (Dt) (48 h)	*p*T (Cv) (96 h)	ECOSAR for *D. tertiolecta*		ECOSAR for *C. vulgaris*	
				Prediction	ECOSAR Class	Prediction	ECOSAR Class
P	Polar narcosis	-0.34	-0.60	0.09	Neutral Organic	0.09	Neutral Organic
2-CP	Polar narcosis	0.24	0.17	0.56	Neutral Organic	0.56	Neutral Organic
3-CP	Polar narcosis	0.68	0.36	0.81	Neutral Organic	0.81	Neutral Organic
4-CP	Polar narcosis	0.41	0.46	0.73	Neutral Organic	0.73	Neutral Organic
2,3-DCP	Polar narcosis	0.85	1.09	1.04	Neutral Organic	1.04	Neutral Organic
2,4-DCP	Polar narcosis	1.02	1.24	1.19	Neutral Organic	1.19	Neutral Organic
2,5-DCP	Polar narcosis	0.97	1.12	1.19	Neutral Organic	1.19	Neutral Organic
2,6-DCP	Polar narcosis	0.19	0.88	0.98	Neutral Organic	0.98	Neutral Organic
3,4-DCP	Polar narcosis	1.29	1.47	1.38	Neutral Organic	1.38	Neutral Organic
3,5-DCP	Polar narcosis	1.42	1.67	1.58	Neutral Organic	1.58	Neutral Organic
2,3,4-TCP	Polar narcosis	1.53	1.64	1.71	Neutral Organic	1.71	Neutral Organic
2,3,5-TCP	Polar narcosis	1.56	1.86	1.73	Neutral Organic	1.73	Neutral Organic
2,3,6-TCP	Polar narcosis	0.67	1.51	1.68	Neutral Organic	1.68	Neutral Organic
2,4,5-TCP	Polar narcosis	1.82	1.67	1.65	Neutral Organic	1.65	Neutral Organic
2,4,6-TCP	Polar narcosis	0.66	1.53	1.63	Neutral Organic	1.63	Neutral Organic
3,4,5-TCP	Polar narcosis	2.18	2.18	1.85	Neutral Organic	1.85	Neutral Organic
2,3,4,5-TCP	Uncoupling	2.32	2.34	1.99	Neutral Organic	1.99	Neutral Organic
2,3,4,6-TCP	Uncoupling	1.92	1.45	2.15	Neutral Organic	2.15	Neutral Organic
2,3,5,6-TCP	Uncoupling	1.99	1.43	1.93	Neutral Organic	1.93	Neutral Organic
PCP	Uncoupling	2.88	1.45	2.62	Neutral Organic	2.62	Neutral Organic
Pyrogallol	Electrophilic	0.47	1.28	1.33	Polyphenols	1.33	Polyphenols
Hq	Electrophilic	2.51	0.02	2.41	Hydroquinones	-0.52	Neutral organic
ClHq	Electrophilic	2.53	1.13	2.62	Hydroquinones	0.04	Neutral organic
TetraClHq	Electrophilic	1.64	1.46	1.98	Neutral Organic	1.98	Neutral Organic
Cat	Electrophilic	0.58	0.27	1.30	Polyphenols	1.30	Polyphenols
4ClCat	Electrophilic	0.73	1.13	1.57	Polyphenols	1.57	Polyphenols
35DiClCat	Electrophilic	1.00	1.95	1.86	Polyphenols	1.86	Polyphenols
Res	Polar narcosis	-0.48	-0.49	-0.37	Neutral Organic	-0.37	Neutral Organic
4ClRes	Polar narcosis	0.02	0.27	0.32	Neutral Organic	0.32	Neutral Organic
46DiClres	Polar narcosis	0.49	1.02	0.76	Neutral Organic	0.76	Neutral Organic

For reactive chemicals, and particularly for hydroquinones, it is clear that the response by *D. tertiolecta* and *C. vulgaris* was different towards these chemicals. For marine algae, the ECOSAR class “hydroquinones” provided the best predictions except for the toxicity of

TetraClHq, which was better predicted by the neutral organic SAR. This is likely to be a consequence of the high lipophilicity of this compound which resulted in the convergence of its MOA to polar narcosis. For freshwater alga *C. vulgaris*, the toxicity of hydroquinones were highly overestimated by the hydroquinone class SAR. The use of neutral organic class for the two hydroquinones resulted in a better prediction of *C. vulgaris* toxicity. However, for hydroquinone and chlorohydroquinone, the use of neutral organic SAR resulted in the under-estimation of toxicity by approximately 0.50 and 1.00 log units, respectively. When results for both algal data sets were evaluated together, the toxicity of hydroquinone differed by more than 200-fold between marine and freshwater algae, the former being much more sensitive. It should be noted that previously Devillers et al. (1990) reported the algal toxicity of hydroquinone to be 0.33 mg/L to freshwater alga *P. subcapitata*, which is very close to the one reported for marine alga *D. tertiolecta*. This implies that for hydroquinones, the marine or freshwater distinction might not be relevant and the response to this class of compounds is most likely a result of species specific phenomenon. Therefore, for regulatory usage of the ECOSAR predictions, the use of 'hydroquinone class' which provided the most conservative prediction class for hydroquinones is recommended to be protective of aquatic ecosystems. As for catechols and pyrogallol, the polyphenol class provided acceptable predictions. This suggests that for *ortho*-dihydroxybenzenes, the use of polyphenols class seems suitable for algae.

The relationship between marine algal toxicity and ECOSAR predictions, as performed using the ECOSAR classes in Table 4.23, is illustrated in Figure 4.20.

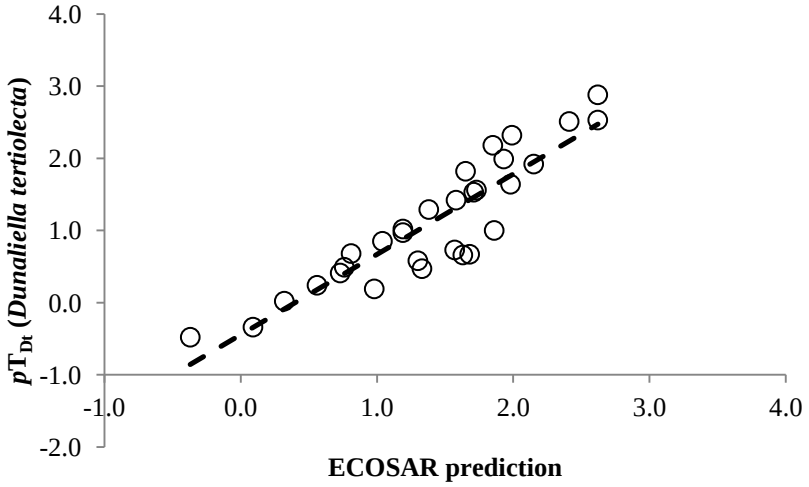


Figure 4.20. ECOSAR predictions for the toxicity of phenols to *D. tertiolecta*

As can be seen from the visual analysis of the ECOSAR predictions using marine algal toxicity data, the prediction was highly significant and there was no outlier (Eq. 4.15):

$$p\mathrm{T}_{\mathrm{Dt}} = 1.11\ (p\mathrm{T}_{\mathrm{ECOSAR}}) - 0.44 \qquad (4.15)$$

(n=30, r^2=0.82, r^2_{cv}= 0.80, *RMSE*=0.36)

The error of the model represented by Eq. 4.15 is also relatively low, implying reliable prediction by the ECOSAR. The predictive performance of ECOSAR for *C. vulgaris* toxicity, on the other hand, was not very high (Eq. 4.16, Figure 4.21).

$$p\mathrm{T}_{\mathrm{Cv}} = 0.79\ (p\mathrm{T}_{\mathrm{ECOSAR}}) + 0.13 \qquad (4.16)$$

(n=30, r^2=0.67, r^2_{cv}= 0.60, *RMSE*=0.42)

Investigation of the ECOSAR predictions revealed that while the toxicity of PCP was overestimated by the model, the toxicity of ClHq was severely underestimated (Table 4.23). Removal of these compounds from the analysis significantly improved the relationship:

$$p\mathrm{T}_{\mathrm{Cv}} = 0.95\ (p\mathrm{T}_{\mathrm{ECOSAR}}) - 0.08 \qquad (4.17)$$

(n=28, r^2=0.78, r^2_{cv}= 0.74, *RMSE*=0.37)

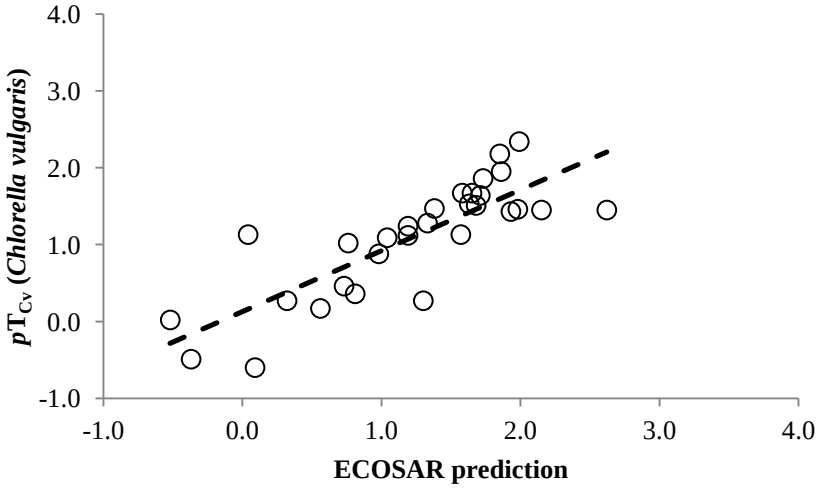


Figure 4.21. ECOSAR predictions for the toxicity of phenols to *C. vulgaris*

In conclusion, the use of ECOSAR to model marine algal toxicity of phenols to *D. tertiolecta* was found to result in acceptable predictions. On the other hand, ECOSAR failed to predict with confidence the toxicity of PCP and ClHq to *C. vulgaris,* resulting in overestimation and underestimation of toxicity, respectively. From a risk assessment perspective, the underestimation of the ClHq toxicity (the toxicity of hydroquinone was also underestimated by 0.50 log units), suggests that the use of ECOSAR in predicting the toxicity of compounds with hydroquinone structure should be executed with care (for instance by applying additional safety factors) to be protective of *C. vulgaris*.

4.4.2. ECOSAR classification performance

Classification and labeling involves an evaluation of the intrinsic hazard of a chemical. This evaluation must be made as set out in the new Classification, Labeling and Packaging Regulation (EC, 2008 - CLP Regulation) for any substance or mixture/preparation manufactured in or imported for the EU. Currently there are more than 7000 hazardous substances listed in the Annex VI to the CLP Regulation however, the number of hazardous chemicals used in EU market is much bigger. By January 3rd 2011, European Chemicals Agency, ECHA had received 3114835 notifications of 24529 substances for the Classification and Labeling Inventory (Aruoja et al., 2011). Therefore, a system that correctly classifies the

toxicity of chemicals is of utmost importance. To this end, the classification ability of ECOSAR predictions was assessed using the algal toxicity data generated in his study.

Figure 4.22 illustrates the toxic class of each tested phenol in *D. tertiolecta* and *C. vulgaris* test system.

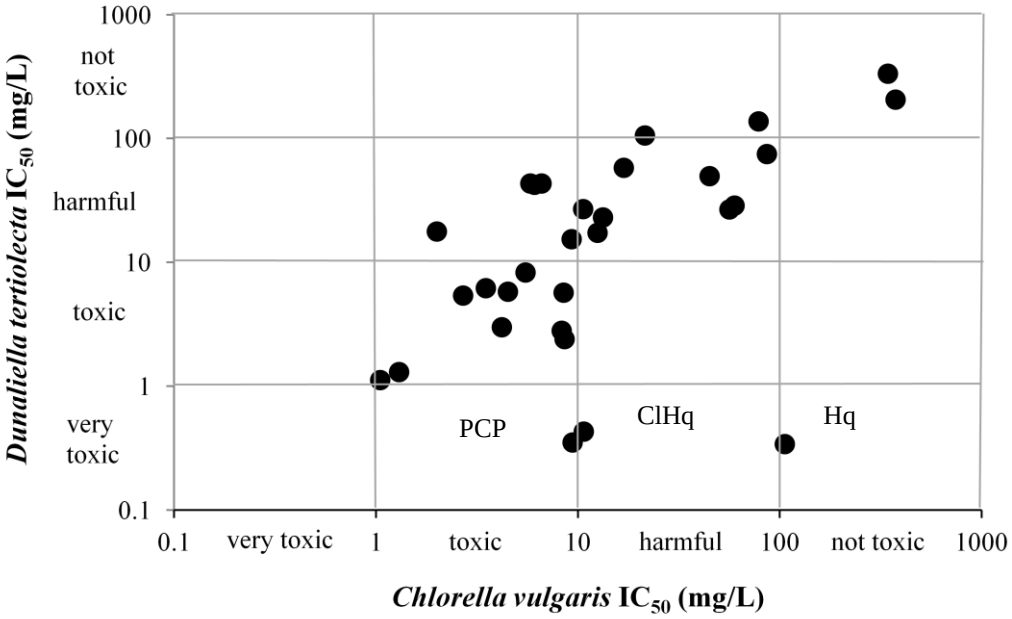


Figure 4.22. Toxic class of phenols towards *D. tertiolecta* and *C. vulgaris*

Comparing the classification of phenols based on the ECOSAR predictions, it was found that 19 compounds were classified correctly for *D. tertiolecta* (Table 4.24). However, in all the misclassified cases, the ECOSAR over-estimated the toxic class for these chemicals which is a better case than under-estimation. As for *C. vulgaris*, ECOSAR classification was correct for 24 chemicals out of 30. The misclassified six compounds were assigned a more toxic class than observed. Similar to the case of marine alga this is not problematic from a risk assessment perspective.

Table 4.24. ECOSAR classification of the algal toxicity of phenols

Compound	Toxic class		ECOSAR classification	
	D. tertiolecta	***C. vulgaris***	***D. tertiolecta***	***C. vulgaris***
P	Not classified	Not classified	Harmful	Harmful
2-CP	Harmful	Harmful	Harmful	Harmful
3-CP	Harmful	Harmful	Harmful	Harmful
4-CP	Harmful	Harmful	Harmful	Harmful
2,3-DCP	Harmful	Harmful	Harmful	Harmful
2,4-DCP	Harmful	Toxic	Toxic	Toxic
2,5-DCP	Harmful	Harmful	Toxic	Toxic
2,6-DCP	Not classified	Harmful	Harmful	Harmful
3,4-DCP	Toxic	Toxic	Toxic	Toxic
3,5-DCP	Toxic	Toxic	Toxic	Toxic
2,3,4-TCP	Toxic	Toxic	Toxic	Toxic
2,3,5-TCP	Toxic	Toxic	Toxic	Toxic
2,3,6-TCP	Harmful	Toxic	Toxic	Toxic
2,4,5-TCP	Toxic	Toxic	Toxic	Toxic
2,4,6-TCP	Harmful	Toxic	Toxic	Toxic
3,4,5-TCP	Toxic	Toxic	Toxic	Toxic
2,3,4,5-TCP	Toxic	Toxic	Toxic	Toxic
2,3,4,6-TCP	Toxic	Toxic	Toxic	Toxic
2,3,5,6-TCP	Toxic	Toxic	Toxic	Toxic
PCP	Very toxic	Toxic	Very toxic	Very toxic
Pyrogallol	Harmful	Toxic	Toxic	Toxic
Hq	Very toxic	Not classified	Very toxic	Not classified
ClHq	Very toxic	Harmful	Very toxic	Toxic
TetraClHq	Toxic	Toxic	Toxic	Toxic
Cat	Harmful	Harmful	Toxic	Toxic
4ClCat	Harmful	Harmful	Toxic	Toxic
35DiClCat	Harmful	Toxic	Toxic	Toxic
Res	Not classified	Not classified	Not classified	Not classified
4ClRes	Not classified	Harmful	Harmful	Harmful
46DiClres	Harmful	Harmful	Harmful	Harmful

The ECOSAR program is used to predict the aquatic toxicity of chemicals based on their similarity of structure. Since 1981, the USEPA has used QSTRs to predict the aquatic toxicity of new industrial chemicals. The acute toxicity of a chemical to fish, water fleas and green algae has been the focus of the development of QSTRs because these organisms are considered as model-organisms (Sanderson et al., 2008). Within the scope of this thesis, ECOSAR predictions for algae were tested on the novel toxicity data set of phenols generated in this thesis. When predictions for both algae are considered, the toxicity classification using the ECOSAR predictions were more protective if the response from the more sensitive species was taken into account. For example, *D. tertiolecta* was found to be much more sensitive to hydroquinone than *C. vulgaris*. Consequently, the ECOSAR's "hydroquinone" class provided a much better estimate of the hydroquinone toxicity to marine alga but using the same

modeling class, the toxicity of this compound was highly over-estimated for *C. vulgaris*. Nevertheless, the "hydroquinone" class provides more protective predictions; therefore, the ECOSAR models should be constructed using this class when the toxicity of phenols with hydroquinone structure is to be predicted.

5. CONCLUSION

In this thesis, the toxicity of 30 phenolic compounds was determined in the same laboratory according to the standardized algal growth inhibition tests using marine alga *Dunaliella tertiolecta* and freshwater alga *Chlorella vulgaris* as the test organisms. The analysis of the mechanistic makeup of the dataset revealed that 19 chemicals out of 30 are expected to exhibit algal toxicity by polar narcosis. In addition, four respiratory uncouplers and seven reactive chemicals (i.e., those acting via electrophilic mechanisms of action) were identified. The fact that the great majority of industrial organic pollutants exhibit a narcotic mode of toxic action (approximately 60-70%), and the presence of compounds having different toxicity mechanisms, the novel algal toxicity data set selected for toxicological assessment is considered to be a miniature and realistic representation of industrial chemical space; as such, it provided a suitable basis for the construction of QSTRs for both algae. The QSTRs were developed following the OECD principles which require that models should have (a) a defined endpoint; (b) an unambiguous algorithm; (c) a defined domain of applicability; (d) appropriate measures of goodness-of-fit, robustness and predictivity; and (e) a mechanistic interpretation, if possible.

The analyses revealed that hydrophobicity underpins both toxicity data sets. From a toxicological perspective, partitioning of a hydrophobic compound into membranes causes disturbances in the structure and functioning of the membranes leading to the so-called baseline toxicity, or narcosis, which constitutes the minimal toxicity of any hydrophobic pollutant. Narcosis is the least specific, but arguably the most important mode of toxic action in terms of risk assessment because approximately 60-70% of all organic industrial chemicals are believed to act via narcosis. As such, European Union Technical Guidance Document proposed several equations employing the hydrophobicity parameter, the logarithm of the *n*-octanol/water partition coefficient (Log K_{ow} or Log *P*), for the use of these relationships in aquatic risk assessment. Although there is a benchmark equation to be used in risk assessment for freshwater algae (reported toxicity to: *Pseudokirchneriella subcapitata*) regarding non-

polar narcotics, there is no equation available for polar narcotics. As for marine algae, there is no equation either for non-polar or polar narcotics. In this respect, the hydrophobicity-based models developed for the first time in this thesis for *D. tertiolecta* and *C. vulgaris*, describing the relationship between the algal toxicity of polar narcotic phenols and Log *P* makes an important contribution to risk assessment of chemicals to algae.

Although hydrophobicity successfully quantifies the transport of a pollutant through the cellular membranes, it does not offer any insights as to the interactions taking place between the xenobiotic and the cellular components in the site of action. As such, it was observed that hydrophobicity alone was not able to quantify the toxicity of all tested phenols to algae in this thesis. Some compounds, particularly polyphenols in which the hydroxyl groups are oriented *ortho* (i.e., catechols and pyrogallol) and *para* (i.e., hydroquinones) to one another, which are potentially capable of being oxidized to reactive species, displayed algal toxicity in excess of that predicted by hydrophobicity. On the other hand, in line with the mechanistic organic chemistry principles, polyphenols in which the hydroxyl groups are oriented *meta* (i.e., resorcinols) to one another were found to act by polar narcosis.

The so-called ‘reactive chemicals’ are responsible for a wide variety of adverse outcomes and modeling these adverse outcomes has been regarded as a serious challenge in the literature because many chemicals are ‘‘activated’’ by metabolic processes and their toxicity might be due to the fact that aromatic compounds are hydroxylated by intracellular mono-oxygenases *in vivo*. In the QSTR models developed for each alga, the presence of reactive toxicants necessitated the selection of a descriptor to account for the variance inflicted by these toxicants. As expected, these descriptors were found to be related to the electrophilicity of the compounds.

In the marine algal QSTRs, the 3D-MoRSE descriptors - particularly the one weighted by the atomic polarizability of the phenols - were found to have good correlation to toxicity. The two-parameter global model employing a 3D-MorSE descriptor (Mor24m) and a charge-related descriptor (C_2) produced excellent statistics for the training and test sets. In the QSTRs developed for freshwater alga *C. vulgaris*, a WHIM descriptor weighted by the atomic mass of

phenols (Tm) and a descriptor which quantifies the number of intra-molecular hydrogen bonds in polyphenols with *ortho* OH substitution (nHBonds) also resulted in a quality QSTR.

The QSTR model developed using the toxicity of phenols to marine alga *D. tertiolecta* as the dependent variable and toxicity data on freshwater alga *C. vulgaris* as the independent variable revealed an important distinction between toxicity data obtained from two distinct test systems. A descriptor responsible for quantifying the polarizability of phenols was required to explain marine algal toxicity of phenols using data on freshwater algae. This might be because the rates of many organic reactions, including those between electrophiles and nucleophiles, are significantly influenced by the polarity of the reaction medium.

One of the novelties and strengths of the QSTR models developed throughout this thesis was that they were further validated externally using a toxicity data set of 58 compounds, including phenols and anilines, towards freshwater alga *P. subcapitata,* which is a freshwater algal species conducted routinely in ecotoxicity testing. The external data set was generated in the same laboratory, by the same researcher following standardized algal growth inhibition test protocols; therefore, it was selected specifically to evaluate the external predictivity of the algal QSTRs constructed in this thesis. Accordingly, the QSTRs developed in this thesis not only predicted the toxicity of phenols to *D. tertiolecta* and *C. vulgaris*, but they displayed excellent external predictivity for *P. subcapitata* as well. In this respect, the models developed in this thesis were shown to be applicable to compounds other than phenols (i.e., anilines) and at least to another algal species, *P. subcapitata*. The models should be further tested in their predictivity using a toxicity data set for other algal species, when such a data set will be available in the literature.

For modeling practice, the comparison between the performances of linear and non-linear models suggested that there was not a big difference in their predictivity as the results highlighted that for the data set studied the multiple linear regression and neural network models performed almost equally well. Considering the transparency and ease-of-application, the linear models should suffice in modeling algal toxicity of phenols to *D. tertiolecta* and *C. vulgaris*.

The data set presented in this study, as mentioned earlier, was specifically selected to be representative industrial chemical space. However, from a statistical perspective, the dominance in the data set of compounds acting through similar toxic modes of action can be expected to result in the selection of a descriptor, or a set of descriptors, that best explain these mechanisms. On the other hand, compounds acting through other modes of action and consequently the associated descriptors may not be equally represented in the models purely due to issues related to the data set bias. For instance, the data set presented in this thesis include chlorophenols and resorcinols, which act through non-covalent mediated interactions; therefore, the algal QSTRs selected a descriptor to explain the toxicity of these compounds first to be able to account for the dominant part of the variance. The remaining descriptors were likely to be selected in the models mainly to account for the variance in the toxicity of compounds acting through covalent interactions. The dominance of polar narcotics in the data set investigated in this thesis provides an important illustration into issues relating to data set bias. In particular, most chemical inventories, and their associated databases, will be biased to some mode(s) of action, and will not be expected to have equal proportions of all toxic mechanisms. In terms of the development of models for regulatory purposes the mechanistic make up of the data set will influence the final model. Since the data set presented in this thesis is comprised of compounds eliciting toxicity mainly by hydrophobicity dependent mechanisms, the reactivity dependent mechanisms might not be well covered. Therefore, the dataset generated in this study can be extended to include other reactive toxicants. By this means, QSTR models capable of explaining broader toxic mechanisms can be developed.

Risk assessment has traditionally focused on determining the potential adverse effects of chemicals to freshwater ecosystems; consequently, this has led to accumulation of ecotoxicity data on freshwater species while data on organisms living in marine environments, particularly on marine algae, has lagged far behind. Ideally, to assess the potential impact of chemicals entering marine ecosystems, any hazard or risk assessment should be based upon data generated using a range of ecologically relevant marine species such as algae, invertebrates and fish. Once chemicals are released into the environment, they can cause adverse effects in non-target organisms in marine environments. Therefore, information on the ecotoxicity of chemicals to non-target species such as marine algae, especially in countries like Turkey

which is surrounded by sea from three sides, should be of utmost importance. It is of our belief that test results obtained from experiments conducted with marine algae should be incorporated to regulatory schemes in Turkey.

The marine algal toxicity data generated in this thesis using a representative species of marine environments, namely *Dunaliella tertiolecta*, makes an important contribution by providing toxicity of 30 phenols for the first time. However, considering the huge gap in marine ecotoxicity, laboratory testing of chemical is not enough to tackle this problem. From this perspective, one of the most important results obtained in this thesis was that the toxicity of polar narcotic phenols can be extrapolated between different species, including algae, with confidence. This is particularly important considering the huge number of industrial chemicals exhibiting toxicity by narcotic mechanisms. This finding is also very important to use toxicity data on freshwater species *in lieu* of marine data. In this respect, several interspecies toxicity correlations suggested that it was viable to use data on one species to predict the toxicity of compounds to another. However, the toxicity of reactive toxicants, particularly hydroquinones, was shown to vary considerably between species. Therefore, their toxicity should ideally be determined in the laboratory for new species. In this respect, the hydroquinone toxicity on marine alga *D. tertiolecta* can serve as a means for protecting marine algae and consequently the marine ecosystems.

Within the scope of this thesis, the widely used ECOSAR programme was also tested in its predictivity. Based on the predictions, the use of ECOSAR for marine alga *D. tertiolecta* produced acceptable results while its use in predicting the toxicity of compounds to *C. vulgaris* acting through mode of actions other than polar narcosis produced some unacceptable results, especially for hydroquinone and chlorohydroquinone. The use of "hydroquinone" class in the ECOSAR predictions is recommended to be protective of aquatic ecosystems for this class of compounds.

In conclusion, this study not only enabled the circulation of new toxicity data on marine and freshwater algae, but also provided detailed assessment on how to integrate it to available scientific knowledge. Based on the knowledge generated in this study, it can be concluded that

the toxicity of polar narcotic compounds to marine alga can be predicted using available toxicity data generated from other test systems. The toxicity models generated throughout the thesis in the form of QSTRs can be used with confidence for compounds within the applicability domain of each proposed QSTR. When available, the global models predictions should be compared to those of local models to ensure reliability in predictions.

REFERENCES

Adema, D.M.M., Vink, G.J., 1981. A comparative study of the toxicity of 1,1,2-trichloroethane, dieldrin, pentachlorophenol and 3,4-dichloroaniline for marine and fresh water organisms. Chemosphere, 10, 6, 533-554.

APHA, AWWA, WEF, 1998. American Public Health Association, American Water Works Association, and Water Environment Federation, Standard Methods for the Examination of Water and Wastewater, 20th Ed. Washington, DC. USA.

Aptula, A.O., Jeliazkova, N.G., Schultz, T.W., Cronin, M.T.D., 2005a. The better predictive model: high Q^2 for the training set or low root mean square error of prediction for the test tet? QSAR & Combinatorial Science, 24, 385-396.

Aptula, A.O., Roberts, D.W., Cronin, M.T.D., Schultz, T.W., 2005b. Chemistry-toxicity relationships for the effects of di- and trihydroxybenzenes to *Tetrahymena pyriformis*. Chemical Research in Toxicology, 18, 844-854.

Aptula A.O., Roberts, D.W., Cronin, M.T.D., 2005c. From experiment to theory: molecular orbital parameters to interpret the skin sensitization potential of 5-chloro-2-methylisothiazol-3-one and 2-methylisothiazol-3-one. Chemical Research in Toxicology, 18, 324-329.

Aptula, A.O., Roberts, D.W., 2006. Mechanistic applicability domains for nonanimal-based prediction of toxicological end points: general principles and application to reactive toxicity. Chemical Research in Toxicology, 19, 1097-1105.

Aptula, A.O., Enoch, S.J., Roberts, D.W., 2009. Chemical mechanisms for skin sensitization by aromatic compounds with hydroxy and amino groups. Chemical Research in Toxicology, 22, 1541-1547.

Aruoja, V., Mariliis, S., Henri-Charles, D., Anne, K., 2011. Toxicity of 58 substituted anilines and phenols to algae *Pseudokirchneriella subcapitata* and bacteria *Vibrio fischeri*: comparison with published data and QSARs. Chemosphere, 10, 1310-1320.

Bajot, F., Cronin, M.T.D., Roberts, D.W., Schultz, T. W., 2011. Reactivity and aquatic toxicity of aromatic compounds transformable to quinone-type Michael acceptors. SAR and QSAR in Environmental Research, 22, 51-65.

Baskin, I.I., Palyulin, V.A., Zefirov, N.S., 2008. Neural networks in building QSAR models. In: Artificial Neural Networks, Methods and Applications, Livingstone, D.J. (ed.), Humana press, Totowa, NJ, pp. 137-158.

Beijerinck, M.W., 1890. Culturversuche mit Zoochlorellen, Lichenengonidien und anderen niederen Algen I-III. Bot. Ztg., 48: 726–740; cited from: Darienko, T., Lydia G., Opayi M., Cecilia R. M., Rhena S., Ulf K., Friedl, T., Pröschold, T., 2010. Chloroidium, a common terrestrial coccoid green alga previously assigned to Chlorella (Trebouxiophyceae, Chlorophyta). European Journal of Phycology, 45, 79-95.

Bhhatarai, B., Gramatica, P., 2010. Per- and polyfluoro toxicity (LC(50) inhalation) study in rat and mouse using QSAR modeling. Chemical Research in Toxicology, 23, 528-539.

Billoir, E., Péry, A.R.R., Charles, S., 2007. Integrating the lethal and sublethal effects of toxic compounds into population dynamics of *Daphnia magna*: a combination of the DEBtox and matrix population models. Ecological Modelling, 203, 204-214.

Borgert, C.J., Quill, T.F., McCarty, L.S., Mason, A.M., 2004. Can mode of action predict mixture toxicity for risk assessment? Toxicology and Applied Pharmacology, 201, 85– 96.

Boyce, D.G., Lewis, M.R., Worm, B., 2010. Global phytoplankton decline over the past century. Nature, 466, 591-596.

Boyd, E.M., Killham, K., Meharg, A.A., 2001. Toxicity of mono-, di- and tri-chlorophenols to lux marked terrestrial bacteria, Burkholderia species Rasc c2 and *Pseudomonas fluorescens*. Chemosphere, 43, 157-166.

Bradbury, S.P., Lipnick, R.L., 1990. Introduction: structural properties for determining mechanisms of toxic action. Environmental Health Perspectives, 87, 181-182.

Breitholtz, M., Ruden, C., Hansson, S.O., Bengtsson, B-E., 2006. Ten challenges for improved ecotoxicological testing in environmental risk assessment. Ecotoxicology and Environmental Safety, 63, 324–335.

Briens, F., Bureau, R., Rault, S., 1999. Applicability of CATALYST in ecotoxicology, a new promising tool for 3D-QSAR: study of chlorophenols. Ecotoxicology and Environmental Safety, 43, 241-51.

Cai, X.Y., Ye, J., Sheng, G., Liu, W., 2009. Time-dependent degradation and toxicity of diclofop-methyl in algal suspensions. Environmental Science and Pollution Research, 16, 459-465.

CCME, 2007. Canadian Council of Ministers of the Environment, Canadian Water Quality Guidelines for the Protection of Aquatic Life, A protocol for the derivation of water quality guidelines for the protection of aquatic life 2007. Ottawa, ON, Canada.

Chon, T-S., 2011. Self-organizing maps applied to ecological sciences, Ecological Informatics, 6, 50-61.

Christensen, E.R., Kusk O.K., Nyholm, N., 2009. Dose-response regressions for algal growth and similar continuous endpoints: calculation of effective concentrations. Environmental Toxicology and Chemistry, 28, 4, 826-835.

Comber, M.H.I., Walker, J.D., Watts, C., Hermens, J., 2003. Quantitative struncture–activity relationships for predicting potential ecological hazard of organic chemicals for use in regulatory risk assessments. Environmental Toxicology and Chemistry, 22, 8, 1822–1828.

Consonni, V., Todeschini, R., Pavan, M., 2002. Structure/Response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. Journal of Chemical Information and Computer Sciences, 42, 682-692.

Consonni, V., Ballabio, D., Todeschini, R., 2009. Comments on the definition of the Q^2 parameter for QSAR validation. Journal of Chemical Information and Modeling, 49, 1669-1678.

Consonni, V., Todeschini, R., 2010. Molecular descriptors. In: Recent Advances in QSAR studies. Puzyn, T., Leszczynski, J., Cronin, M.T.D. (Eds), Springer, New York.

Consonni, V., Todeschini, R. 2011. Multivariate analysis of molecular descriptors. In: Statistical Modelling of Molecular Descriptors in QSAR/QSPR, Wiley-VCH, Weinheim.

Costanza, R., Hartung, T., 2009. Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements for chemicals–a report by the Transatlantic Think Tank for Toxicology. ALTEX, 26, 187–208.

Crisinel, A., Delauny, L., Rossel, D., Tarradellas, J., Meyer, H., Saïah, H., Vogel, P., Delisle, C., Blaise, C., 1994. Cyst-based ecotoxicological tests using anostracans: comparison of two species of *Streptocephalus*. Environmental Toxicology and Water Quality, 9, 317-326.

Cronin, M.T.D., Schultz, T.W., 1998. Structure–toxicity relationships for three mechanisms of action of toxicity to *Vibrio fischeri*. Ecotoxicology and Environmental Safety, 39, 65-69.

Cronin, M.T.D., Aptula, A.O., Duffy, J.C., Netzeva, T.I., Rowe, P.H., Valkova, I.V., Schultz, T.W., 2002. Comparative assessment of methods to develop QSARs for the prediction of the toxicity of phenols to *Tetrahymena pyriformis*. Chemosphere, 49, 1201-1221.

Cronin, M.T.D., Schultz, T.W., 2003. Pitfalls in QSAR. Journal of Molecular Structure, 622, 39-51.

Cronin, M.T.D., Netzeva, T.I., Dearden, J.C., Edwards, R., Worgan, A.D.P., 2004. Assessment and modeling of the toxicity of organic chemicals to *Chlorella vulgaris*: development of a novel database. Chemical Research in Toxicology, 17, 545-554.

Cros, A. F. A., PhD Thesis, University of Strasbourg 1863; cited from: Kubinyi, H., 2002. From Narcosis to Hyperspace: The History of QSAR, Quantitative Structure-Activity Relationships, 21, 348-356.

Czaplicka M., 2004. Sources and transformations of chlorophenols in the natural environment. Science of the Total Environment, 322, 21-39.
Darienko, T., Lydia G., Opayi M., Cecilia R.M., Rhena S., Ulf, K., Friedl, T., Pröschold, T., 2010. Chloroidium, a common terrestrial coccoid green alga previously assigned to Chlorella (Trebouxiophyceae, Chlorophyta). European Journal of Phycology, 45, 79-95.

DeLorenzo, M. E., Serrano, L. 2006. Mixture toxicity of the antifouling compound irgarol to the marine phytoplankton species *Dunaliella tertiolecta*. Journal of Environmental Science and Health, Part B, Pesticides, Food Contaminants, and Agricultural Wastes, 41, 1349-1360.

DeLorenzo, M.E., Fleming, J., 2008. Individual and mixture effects of selected pharmaceuticals and personal care products on the marine phytoplankton species *Dunaliella tertiolecta*. Archives of Environmental Contamination and Toxicology, 54, 203-210.

DeLorenzo, M.E., 2009. Utility of *Dunaliella* in ecotoxicity testing. In: The Alga *Dunaliella*: Biodiversity, Physiology, Genomics and Biotechnology, A. Ben-Amotz, J.E.W. Polle, D.V.S. Rao, Eds., Science Publishers, Enfield, NH, pp. 495-512.

Devillers, J., Boule, P., Vasseur, P., Prevot, P., Steiman, R., Seigle-Murandi, F., Benoit-Guyod, J.L., Nendza, M., Grioni, C., Dive, D., Chambon, P., 1990. Environmental health risks of hydroquinone. Ecotoxicology and Environmental Safety, 19, 327-354.

Devillers, J., 2008. Artificial neural network modeling in environmental toxicology. In: Artificial Neural Networks, Methods and Applications, Livingstone, D.J. (ed.), Humana press, Totowa, NJ, pp. 61-79.

Dimitrov, S.D., Mekenyan, O.G., Schultz, T.W., 2000. Interspecies modeling of narcotics toxicity to aquatic animals. Bulletin of Environmental Contamination and Toxicology, 65, 399–406.

Dimitrov, S.D., Mekenyan, O.G., Sinks, G.D., Schultz, T.W., 2003. Global modeling of narcotic chemicals: ciliate and fish toxicity. Journal of Molecular Structure, 622, 63-70.

Duchowicz, P., Mercader, A.G., Fernandez, F.M., Castro, E.A., 2008. Prediction of aqueous toxicity for heterogeneous phenol derivatives by QSAR. Chemometrics and Intelligent Laboratory Systems, 90, 2, 97-107.

Eberius, M., Guido M., Ilka R., Jörg V., 2002. Sensitivity of different growth inhibition tests-just a question of mathematical calculation? Theory and Practice for Algae and Duckweed. Ecotoxicology, 11, 293-297.

EC, 2006. European Commission, Regulation No. 1907/2006 of the European Parliament and of the Council of 18 December 2006 Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union, L 396/1-849, Brussels, Belgium.

EC, 2008. European Commission, Regulation No. 1272/2008 of the European Parliament and of the Council of 16 December 2008 on Classification, Labelling and Packaging of Substances and Mixtures, Amending and Repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No. 1907/2006. Official Journal of the European Union, L 353, Brussels, Belgium.

ECETOC, 1993. European Centre for Ecotoxicology and Toxicology of Chemicals, Aquatic toxicity data evaluation. Technical Report, No. 56, Brussels, Belgium.

ECETOC, 2001. European Centre for Ecotoxicology and Toxicology of Chemicals, Risk assessment in marine environments. Technical Report, No. 82, Brussels-Belgium.

ECETOC, 2003. European Centre for Ecotoxicology and Toxicology of Chemicals, (Q)SARs: evaluation of the commercially available software for human health and environmental endpoints with respect to chemical management applications. Technical report No. 89, Brussels, Belgium.

ECETOC, 2008. European Centre for Ecotoxicology and Toxicology of Chemicals, Workshop on the probabilistic approaches for marine hazard assessment. Workshop Report No. 15. Brussels, Belgium.

ECOSAR, 2011. ECOlogical Structure-Activity Relationship Model, Methodology Document, Estimating toxicity of industrial chemicals to aquatic organisms using ECOSAR class program. US EPA, Washington, DC.

ECOTOX, 2012. Ecological Toxicity Database, United States Environmental Protection Agency, (http://cfpub.epa.gov/ecotox/) Last Access: February, 2012.

Enoch, S.J., Cronin, M.T.D., Schultz, T.W., Madden, J.C., 2008. An evaluation of global QSAR models for the prediction of the toxicity of phenols to *Tetrahymena pyriformis*. Chemosphere, 71, 1225-1232.

Eroğlu, E., Palaz, S., Oltulu, O., Türkmen, H., Özaydın, C., 2007. Comparative QSTR study using semi-empirical and first principle methods based descriptors for acute toxicity of diverse organic compounds to the fathead minnow. Internation Journal of Molecular Sciences, 8, 1265-1283.

Ertürk, M.D., Saçan, M.T., 2012. First toxicity data of chlorophenols on marine alga *Dunaliella tertiolecta*: correlation of marine algal toxicity with hydrophobicity and interspecies toxicity relationships. Environmental Toxicology and Chemistry, 31, 1113-1120.

Ertürk, M.D., Saçan, M.T., Novic, M., Minovski, N., 2012. Quantitative structure-activity relationships (QSARs) using the novel marine algal toxicity data of phenols. Journal of Molecular Graphics and Modelling, 38, 90-100.

Escher, B.I., Snozzi, M., Schwarzenbach, R.P., 1996. Uptake, speciation, and uncoupling activity of substituted phenols in energy transducing membranes. Environmental Science and Technology, 30, 3071-3079.

Escher, B.I., Hunziker, R., Westall, J.C., Schwarzenbach, R.P., 1999. Kinetic model to describe the intrinsic uncoupling activity of substituted phenols in energy transducing membranes. Environmental Science and Technology, 33, 560–570.

Escher, B.I., Schwarzenbach, R.P., Westall, J.W.C., 2000a. Evaluation of liposome-water partitioning of organic acids and bases: I. Development of sorption model. Environmental Science and Technology, 34, 3954–3961.

Escher, B.I., Schwarzenbach, R.P., Westall, J.C., 2000b. Evaluation of liposome-water partitioning of organic acids and bases: II. Comparison of experimental determination methods. Environmental Science and Technology, 34, 3962–3698.

Escher, B.I., Schwarzenbach R.P., 2002. Mechanistic studies on baseline toxicity and uncoupling of organic compounds as a basis for modeling effective membrane concentrations in aquatic organisms, Aquat. Sci., 64, 20-35.

Escher, B.I., Hermens, J.L.M., 2004. Internal exposure: Linking bioavailability to effects. Environmental Science and Technology, 38, 455A-462A.

EU TGD, 2003. European Union Technical Guidance Document, Risk assessment in support of the commission directive 93/67/EEC on risk assessment for new notified substances Commission Regulation (EC) No 1488/94 on risk assessment for existing substances directive 98/8/EC of the European Parliament and of the council concerning the placing of biocidal products on the market. Environmental Chemicals Bureau, Ispra, Italy.

Faber, K., Kowalski, B.R., 1997. Improved prediction error estimates for multivariate calibration by correcting for the measurement error in the reference values. Applied Spectroscopy, 51, 660–665.

Fahl, G.M., Kreft, L., Altenburger, R., Faust, M., Boedeker, W., Grime, L.H., 1995. *p*H-dependent sorption, bioconcentration and algal toxicity of sulfonylurea herbicides. Aquatic Toxicology, 31, 175-187.

Forbes, V.E., Calow, P., 2002. Extrapolation in ecological risk assessment: balancing pragmatism and precaution in chemical controls legislation. Bioscience, 52, 249–257.

Fujita, T., 1966. The analysis of physiological activity of substituted phenols with substituent constants. Journal of Medicinal Chemistry, 9, 797–803.

Furusjö, E., Svenson, A., Rahmberg, M., Andersson, M., 2006. The importance of outlier detection and training set selection for reliable environmental QSAR predictions. Chemosphere, 63, 99-108.

Geis, S.W., Fleming, K.L., Korthals, E.T., Searle, G., Reynolds, L., Karner, D.A., 2000. Modifications to the algal growth inhibition test for use as a regulatory assay. Environmental Toxicology and Chemistry, 19, 36–41.

Geyer, H., Scheunert, I., Korte, F., 1985. The effects of organic environmental chemicals on the growth of the alga *Scenedesmus subspicatus*; a contribution to environmental biology. Chemosphere, 14, 9, 1355-1369.

Golbraikh, A., Shen, M., Xiao, Z., Xiao, Y.D., Lee, K.H., Tropsha, A., 2003. Rational selection of training and test sets for the development of validated QSAR models. Journal of Computer-aided Molecular Design, 17, 241-53.

Golla, S., Madihally, S., Robinson Jr., R.L., Gasem, K.A.M., 2009. Quantitative structure–property relationships modeling of skin irritation. Toxicology in Vitro, 23, 176-184.

González, M.A., Gómez, P.I., Polle, J.E.W., 2009. Taxonomy and phylogeny of the genus *Dunaliella*. In: The Alga *Dunaliella*: Biodiversity, Physiology, Genomics and Biotechnology, A. Ben-Amotz, J.E.W. Polle, D.V.S. Rao, Eds., Science Publishers, Enfield, NH, pp. 15-44.

Gramatica, P., Giani, E., Papa, E., 2007. Statistical external validation and consensus modeling: a QSPR case study for Koc prediction. Journal of Molecular Graphics and Modelling, 25, 755-766.

Hanrahan, G., 2009. Environmental chemometrics, principles and modern applications. CRC Press, Boca Raton, USA.

Hansch, C., Maloney, P.P., Fujita, T., Muir, R.M., 1962, Nature, 194, 178; cited from: Selassie, C.D., 2003. History of Quantitative Structure-Activity Relationships, In: Burger's Medicinal Chemistry, Drug Discovery and Development, Abraham, D.J. (Ed.), John Wiley&Sons, Inc, 6th edition, vol. 1, pp. 1-47.

Hansch, C., Fujita, T., 1964. ρ-σ-π Analysis. A method for the correlation of biological activity and chemical structure, Journal of American Chemical Society, 86, 1616-1626.

Hansch, C., Leo, A., 1995. Exploring QSAR: fundamentals and applications in chemistry and biology. Americal Chemical Society, Washington, DC.

Heicht-Nielsen, R., 1987. Counter propagation networks. Applied Optics, 26, 4979-4984

Henegar, A., Mombelli, E., Pandard, P., Péry, A.R.R., 2011. What can be learnt from an ecotoxicity database in the framework of the REACH regulation? Science of the Total Environment, 409, 489–494.

Hewitt, M., Cronin, M.T.D., Madden, J.C., Rowe, P.H., Johnson, C., Obi, A., Enoch, S.J., 2007. Consensus QSAR models: Do benefits outweigh the complexity? Journal of Chemical Information and Modeling, 47, 1460-1468.

Hewitt, M., Cronin, M.T.D., Rowe, P.H., Schultz, T.W., 2011. Repeatability analysis of the *Tetrahymena pyriformis* population growth impairment assay. SAR and QSAR in Environmental Research, 22, 621-637.

Hiller, S.A., Glaz, A.B., Rastrigin, L.A., Rosenblit, A.B., 1971. Recognition of physiological activity of chemical compounds on perceptron with random adaptation of structure. Dokl Akad Nauk SSSR, 199, 851-853.

Hoaglin, D.C., Welsch, R.E., 1978. The hat matrix in regression and ANOVA. The American Statistician, 32, 17-22.

Hogl, E., 1958. Some non-volatile extracts of coffee. Mitteilungen aus dem Gebiete der Lebensmitteluntersuchung und Hygiene, 49, 433–441.

Hommen, U., Baveco, J.M.H., Galic, N., van den Brink, P.J., 2010. Potential application of ecological models in the European environmental risk assessment of chemicals. I. Review of protection goals in EU directives and regulations. Integrated Environmental Assessment and Management, 6, 325-337.

Hutchinson, T.H., Scholz, N., Guhl, W., 1998. Analysis of ECETOC aquatic toxicity (EAT) database IV- Comparative toxicity of chemical substances to freshwater versus saltwater organisms. Chemosphere, 36, 1, 143-153.

ISO, 2004. International Organization for Standardization, 8692, Water quality — Freshwater algal growth inhibition test with unicellular green algae". Geneva, Switzerland.

Jensen, J., 1996. Chlorophenols in the terrestrial environment. Reviews of Environmental Contamination and Toxicology, 146, 25-51.

Janssen, M., Slenders, P., Tramper, J., Mur, L.R., Wijffels, R.H., 2001. Photosynthetic efficiency of *Dunaliella tertiolecta* under short light/dark cycles. Enzyme and Microbial Technology, 29, 298 -305.

Jover, J., Bosque, R., Sales, J., 2009. Quantitative structure-property relationship estimation of cation binding affinity of the common amino acids. Journal of Physical Chemistry A**,** 113, 3703–3708.

Kaiser K.L.E., Palabrica V.S., 1991. *Photobacterium phosphoreum* toxicity data index. Water Pollution Research Journal of Canada, 26, 361-431.

Kaiser, K.L.E., 1998. Correlations of *Vibrio fischeri* bacteria test data with bioassay data for other organisms. Environmental Health Perspectives, 106, 583-591.

Kar, S., Roy, K., 2010. First report on interspecies quantitative correlation of ecotoxicity of pharmaceuticals. Chemosphere, 81, 738-747.

Karelson, M., Lobanov, V.S., Katritzky, A.R., 1996. Quantum-chemical descriptors in QSAR/QSPR studies. Chemical Reviews, 96, 1027-1043.

Karelson, M., 2000. Molecular descriptors in QSAR/QSPR. Wiley-Interscience, New York. USA.

Katritzky, A.R., Lobanov, V.S. Karelson, M., 1995a. CODESSA, Training manual, Gainsville

Katritzky, A.R., Lobanov, V.S., Karelson, M. 1995b. QSPR: the correlation and quantitative prediction of chemical and physical properties from structure. Chemical Society Reviews, 24, 279-287.

Kier, L.B., Hall, L.H., 1976. Molecular connectivity in chemistry and drug research. Academic Press, New York.

Könemann, H., Musch, A., 1981. Quantitative structure-activity relationship in fish toxicity studies. Part 2: the influence of *p*H on the QSAR of chlorophenols. Toxicology, 19, 223–228.

Kramer, C., Trümper, L., 1986. Quantitative structure-activity relationships for the autotrophic growth inhibition of synchronous *Chlorella vulgaris* suspensions by monosubstituted phenols. Biochemie und Physiologie der Pflanzen, 181, 81–89.

Kubinyi, H., 2002. From narcosis to hyperspace: the history of QSAR, Quantitative Structure-Activity Relationships, 21, 348-356.

Kuhl, A., Lorenzen H., 1964. Handling and culturing of *Chlorella*. In Prescott, D.M. (Ed.). Methods in Cell Physiology. Academic Press, New York-London, pp. 159-187.

Lee, P.Y., Chen, C.Y., 2009. Toxicity and quantitative structure-activity relationships of benzoic acids to *Pseudokirchneriella subcapitata*. Journal of Hazardous Materials, 165, 156-161.

Lukacinova, A., Mojsiz, J., Benacka, R., Lovasova, E., Hijova, E., Nistiar, F., 2007. *Tetrahymena pyriformis* as a valuable unicellular animal model organism for determination of xenobiotics. Proceedings of 27th International Symposium, Industrial Toxicology, Bratislava, Slovakia.

Lyakhovich, V.V., Vavilin, V.A., Zenkov, N.K., Menshchikova, E.B., 2006. Active defense under oxidative stress. The antioxidant responsive element. Biochemistry (Moscow), 71, 962-974.

Madigan, M., Martinko, J. (eds), 2005. Brock Biology of Microorganisms. 11th Ed. Prentice Hall.

Mayer, P., Frickmann, J., Christensen, E.R., Nyholm, N., 1998. Influence of growth conditions on the results obtained in algal toxicity tests. Environmental Toxicology and Chemistry, 17, 6, 1091–1098.

Mercader, A.G., Pomilio, A.B., 2010. QSAR study of flavonoids and biflavonoids as influenza H1N1 virus neuraminidase inhibitors. European Journal of Medicinal Chemistry, 45, 1724-1730.

Minovski, N. 2009. DragCOD v1.0, laboratory for chemometrics. National Institute of Chemistry, Ljubljana

Mirabelli, M.C., Hoppin, J.A., Tolbert, P.E., Herrick, R.F., Gnepp, D.R., Brann, E.A., 2000. Occupational exposure to chlorophenol and the risk of nasal and nasopharyngeal cancers among U.S. men aged 30 to 60. American Journal of Industrial Medicine, 37, 532-541.

Moosus, M., Maran, U., 2011. Quantitative structure–activity relationship analysis of acute toxicity of diverse chemicals to *Daphnia magna* with whole molecule descriptors. SAR and QSAR in Environmental Research, 22, 757-774.

Murkovski, A., Skórska, E. 2010. Effect of (C6H5)3PbCl and (C6H5)3SnCl on delayed luminescence intensity, evolving oxygen and electron transport rate in photosystem II of *Chlorella vulgaris*. Bulletin of Environmental Contamination and Toxicology, 84, 157-160.

Nendza, M., Seydel, J.K., 1988. Quantitative structure-toxicity relationships for ecotoxicologically relevant biotestsystems and chemicals. Chemosphere, 17, 8, 1585-1602.

Netzeva, T.I., Pavan, M., Worth, A.P., 2008. Review of (quantitative) structure–activity relationships for acute aquatic toxicity. QSAR & Combinatorial Science, 27, 1, 77-90.

Neuwoehner, J., Escher, B.I., 2001. The *p*H-dependent toxicity of basic pharmaceuticals in the green algae *Scenedesmus vacuolatus* can be explained with a toxicokinetic ion-trapping model. Aquatic Toxicology, 101, 266-75.

Nichkova, M., Galve, R., Marco, M. P., 2002. Biological monitoring of 2,4,5-trichlorophenol (II): evaluation of an enzyme-linked immunosorbent assay for the analysis of water, urine, and serum samples. Chemical Research in Toxicology, 15, 1371-1379.

NIWA, 1998. Marine algae *Dunaliella tertiolecta* chronic toxicity test protocol (MFE 80205: Appendix 1). Ecotoxicology Lab, New Zealand.

Novic, M., Vracko, M., 2003. Artificial neural networks in molecular structures-property studies. In: Nature inspired methods in chemometrics: genetic algorithms and artificial neural networks Leardi, R. (Ed.), Elsevier, Data Handling in Science and Technology, 23, 231-256.

Nyholm, N., Källqvist, T., 1989. Methods for growth inhibition toxicity tests with freshwater algae. Environmental Toxicology, 8, 689-703.

Okay O, Gaines A. 1996. Toxicity of 2,4-D acid to phytoplankton. Water Research, 30, 688–696.

Okumura, Y., Koyama, J., Takaku, H., Satoh, H., 2001. Influence of organic solvents on the growth of marine microalgae. Archives of Environmental Contamination and Toxicology, 41, 123-128.

Olivier, S, Scragg, A.H., Morrison, J., 2003. The effect of chlorophenols on the growth of *Chlorella* VT-1. Enzyme and Microbial Technology, 32, 837-842.

OECD, 2006. Organization for Economic Cooperation and Development, Guideline 201: Freshwater alga and cyanobacteria growth inhibition test. Paris, France.

OECD, 2007. Organization for Economic Cooperation and Development, Guidance document on the validation of (quantitative) structure-activity relationships [(Q)SAR] models. OECD Environment Health and Safety Publications: Series on Testing and Assessment No. 69, Paris.

OECD, 2012. Organization for Economic Cooperation and Development, QSAR Application Toolbox, version 2.3, available at: http://www.oecd.org/document/54/0,3746,en_2649_34379_42923638_1_1_1_1,00.html), (Accessed: February, 2013).

Öberg, T., 2004. A QSAR for baseline toxicity: validation, domain of application, and prediction. Chemical Research in Toxicology, 17, 1630-1637.

Papa, E., Villa, F., Gramatica, P., 2005. Statistically validated QSARs, based on theoretical descriptors, for modeling aquatic toxicity of organic chemicals in *Pimephales promelas* (Fathead Minnow). Journal of Chemical Information and Modeling, 45, 1256-1266.

Pepelko, W.E., Gaylor, D.W., Mukerjee, D., 2005. Comparative toxic potency ranking of chlorophenols. Toxicology and Industrial Health, 21, 93-111.

Peters, C., Ahlf, W., von Lochow, H.E.C., Ratte, H.T., 2005. How should we deal with the uncertainty in the extrapolation of the sensitivity of marine organisms to narcotics? Critical review of the state of the art of marine risk assessment and recommendations for future research. MarSens CEFIC LRI-ECO3A-TUHH-0407.

Polle, E.W., Tran, D., Ben-Amotz, A. 2009. History, distribution, and habitats of algae of the genus Dunaliella TEODEROSCO (Chlorophyceae). In: The Alga Dunaliella: Biodiversity, Physiology, Genomics and Biotechnology, A. Ben-Amotz, J.E.W. Polle, D.V.S. Rao, Eds., Science Publishers, Enfield, NH, pp. 1-13.

Raimondo, S., Mineau, P., Barron, M.G., 2007. Estimation of chemical toxicity to wildlife species using interspecies correlation models. Environmental Science and Technology, 41, 5888–5894.

Raimondo, S., Jackson, C.R., Barron, M.G., 2010. Influence of taxonomic relatedness and chemical mode of action in acute interspecies estimation models for aquatic species. Environmental Science and Technology, 44, 7711-7716.

Rao, D.V.S., 2009. Cultivation, growth media, division rates and applications of *Dunaliella* species. In: The Alga Dunaliella: Biodiversity, Physiology, Genomics and Biotechnology, A. Ben-Amotz, J.E.W. Polle, D.V.S. Rao, Eds., Science Publishers, Enfield, NH, pp. 45-89.

Rayne, S., Forest, K., Friesen, K.J., 2009. Mechanistic aspects regarding the direct aqueous environmental photochemistry of phenol and its simple halogenated derivatives. A review. Environment International, 35, 2, 425-437.

Repetto, G., Jos, A., Hazen, M. J., Molero, M. L., del Peso, A., Salguero, M., del Castillo, P., Rodriguez-Vicente, M.C., Repetto, M., 2001. A test battery for the ecotoxicological evaluation of pentachlorophenol. Toxicology in Vitro, 15, 503–509.

Reuschenbach, P., Silvani, M., Dammann, M., Warnecke, D., Knacker, T., 2008. ECOSAR model performance with a large test set of industrial chemicals. Chemosphere, 71, 1986-1995.

Riedl, J., Altenburger, R., 2007. Physicochemical substance properties as indicators for unreliable exposure in microplate-based bioassays. Chemosphere, 67, 2210-2220.

Roy, P. P., Leonard, J. T., Roy, K., 2008. Exploring the impact of size of training sets for the development of predictive QSAR models. Chemometrics and Intelligent Laboratory Systems, 90, 31-42.

Roy, P.P., Paul, S., Mitra, I., Roy, K., 2009. On two novel parameters for validation of predictive QSAR models. Molecules, 14, 1660-1701.

Russo, R.C., 2002. Development of marine water quality criteria for the USA. Marine Pollution Bulletin, 45, 84-91.

Russom, C.L., Bradbury, S.P., Broderius, S.J., 1997. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (Pimephales *promelas*). Environmental Toxicology and Chemistry, 16, 948-967.

Saçan, M.T., Balcıoğlu, I.A., 2006. A case study on algal response to raw and treated effluents from an aluminum plating plant and a pharmaceutical plant. Ecotoxicology and Environmental Safety, 64, 234-243.

Saçan, M.T., Novic, M., Ertürk, M.D., Minovski, N., 2013. *In silico* modelling of *in vivo* toxicity data on marine alga, *D. tertiolecta*. In Advances in Mathematical Chemistry, Basak, S., Restrepo, G. (Eds.) In Press.

Sahinkaya, E., Dilek, F.B., 2009. The growth behavior of *Chlorella vulgaris* in the presence of 4-chlorophenol and 2,4-dichlorophenol. Ecotoxicology and Environmental Safety, 72, 781-786.

Sanderson, H., Johnson, D.J., Wilson, C.J., Brain, R.A., Solomon, K.R., 2003. Probabilistic hazard assessment of environmentally occurring pharmaceuticals toxicity to fish, daphnids and algae by ECOSAR screening. Toxicology Letters, 144, 383-395.

Sanderson, H., Dyer, S.D., Nabholz, J.V., 2008. (Q)SAR and extrapolation. In: Extrapolation practice for ecotoxicological effect characterization of chemicals, Solomon, K.R., Brock, T.C.M., de Zwart, D., Dyer, S.D., Posthuma, L., Richards, S.M., Sanderson, H., Sibley, P.K., van den Brink, P.J., (Eds.), CRC Press, Pensacola, Fla., SETAC.

Satoh, T., Saitoh, S., Hosaka, M., Kosaka, K., 2009. Simple ortho- and para-hydroquinones as compounds neuroprotective against oxidative stress in a manner associated with specific transcriptional activation. Biochemical and Biophysical Research Communications, 379, 537-541.

Schäfer, H., Hettler, H., Fritsche, E., Pitzen, G., Röderer, G., Wenzel, A., 1994. Biotests using unicellular algae and ciliates for predicting long-term effects of toxicants. Ecotoxicology and Environmental Safety, 27, 64-81.

Schultz, T. W., 1987. The use of ionization constant (*p*Ka) in selecting models of toxicity in phenols. Ecotoxicology and Environmental Safety, 14, 178–183.

Schultz, T.W., Sinks, G.D., Cronin, M.T.D., 1997. Quinone-induced toxicity to *Tetrahymena*: structure-activity relationships. Aquatic Toxicology, 39, 267-278.

Schultz, T.W., 1999. Structure-toxicity relationships for benzenes evaluated with *Tetrahymena pyriformis*. Chemical Research in Toxicology, 12, 1262-1267.

Schultz, T.W., Carlson, R.E., Cronin, M.T.D., Hermens, J.L.M., Johnson, R., O'Brien, P.J. Roberts, D.W., Siraki, A., Wallace, K.B., Veith, G.B., 2006. A conceptual framework for predicting the toxicity of reactive chemicals: modeling soft electrophilicity. SAR and QSAR in Environmental Research, 17, 413-428.

Schüürmann G., Segner H., Jung K., 1997. Multivariate mode-of-action analysis of acute toxicity of phenols. Aquatic Toxicology, 38, 277-296.

Schweigert, N., Zehnder, A.J.B., Eggen, R.I.L., 2001a. Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. Environmental Microbiology, 3, 2, 81-91.

Schweigert, N., Hunziker, R.W., Escher, B.I., Eggen, R.I.L., 2001b. Acute toxicity of (chloro) catechols and (chloro) catechol-copper combinations in *Escherichia coli* corresponds to their membrane toxicity in vitro. Environmental Toxicology and Chemistry, 20, 2, 239–247.

Schwöbel, J.A.H., Koleva, Y.K., Enoch, S.J., Bajot, F., Hewitt, M., Madden, J.C., Roberts, D.W., Schultz, T.W., Cronin, M.T.D., 2011. Measurements and estimation of electrophilic reactivity for predictive toxicology. Chemical Reviews, 111, 2562-2596.
Scragg, A.H., Spiller, L., Morrison, J., 2003. The effect of 2,4-dichlorophenol on the microalga *Chlorella* VT-1. Enzyme and Microbial Technology, 32, 616-622.

Scragg, A.H., 2006. The effect of phenol on the growth of *Chlorella vulgaris* and *Chlorella VT-1*. Enzyme and Microbial Technology, 39, 796-799.

Selassie, C.D., 2003. History of quantitative structure-activity relationships, In: Burger's Medicinal Chemistry, Drug Discovery and Development, Abraham, D.J. (Ed.), John Wiley&Sons, Inc, 6th edition, vol. 1, pp. 1-47.

Seward, J.R., Hamblen, E.L., Schultz, T.W., 2002. Regression comparisons of *Tetrahymena pyriformis* and *Poecilia reticulata* toxicity. Chemosphere. 47, 93–101.

Shigeoka, T., Sato, Y., Takeda, Y., Yoshida, I., Yamauchi, F., 1988. Acute toxicity of chlorophenols to green algae *Selenastrum capricornutum* and *Chlorella vulgaris* and quantitative structure–activity relationships. Environmental Toxicology and Chemistry, 7, 847–854.

Sponza, D.T., 2002. Incorporation of toxicity tests into the Turkish industrial discharge monitoring systems. Archives of Environmental Contamination and Toxicology, 43, 186-197.

Spycher, S., Pellegrini, E., Gasteiger, J., 2005. Use of structure descriptors to discriminate between modes of toxic action of phenols. Journal of Chemical Information and Modeling, 45, 200–208.

Spycher, S., Smejtek, P., Netzeva, T.I., Escher, B.I., 2008. Toward a class-independent quantitative structure--activity relationship model for uncouplers of oxidative phosphorylation. Chemical Research in Toxicology, 21, 911-927.

Stom, D.I., Roth R., 1981. Some effects of polyphenols on aquatic plants: I. Toxicity of phenols in aquatic plants. Bulletin of Environmental Contamination and Toxicology, 27, 332-337.

Sverdrup, L.E., Fürst, C.S., Weideborg, M., Vik, E.A., Stenersen, J., 2002. Relative sensitivity of one freshwater and two marine acute toxicity tests as determined by testing 30 off-shore E & P chemicals. Water Technology, 46, 311-318.

Tabata, K., 1962. Toxicity of ammonia to aquatic animals with reference to the effect of *p*H and carbon dioxide. Bulletin of Tokai Regional Fisheries Research Laboratory, 34, 67–74.

Tebby, C., Mombelli, E., Pandard, P., Péry, A.R.R., 2011. Exploring an ecotoxicity database with the OECD (Q)SAR Toolbox and DRAGON descriptors in order to prioritise testing on algae, daphnids, and fish. Science of the Total Environment, 409, 3334–3343.

TerraTox™ database, 2006, version 3.0, Terrabase, Inc., Hamilton, Ontario, Canada

TETRATOX™ database, www.vet.utk.edu/TETRATOX (Last accessed: February, 2013)

Todeschini, R., Gramatica, P., 1997. The WHIM theory: new 3D molecular descriptors for QSAR in environmental modeling, SAR and QSAR in Environmental Research, 7, 89-115.

Todeschini, R., Consonni, V., 2000. Handbook of molecular descriptors. Wiley-VCH, Weinheim.

Todeschini, R., Consonni, V., 2009. Molecular descriptors for chemoinformatics. Vol. 41, 2nd, revised and enlarged ed. Wiley-VCH, Weinheim.

Topping, D.C., Bernard, L.G., O'Donoghue, J.L., English, J.C., 2007. Hydroquinone: acute and subchronic toxicity studies with emphasis on neurobehavioral and nephrotoxic effects. Food and Chemical Toxicology, 45, 70–78.

UN, 1992. United Nations, Agenda 21: Programme for action for sustainable development. Rio declaration on environment and development. The final text of agreements negotiated by governments at the UN conference on environment and development, 3-14 June, 1992, Rio de Janeiro, Brazil. E.93.1.11. UN Department of Public Information.

Upadhyay, G., Gupta, S.P., Prakash, O., Singh, M.P., 2010. Pyrogallol-mediated toxicity and natural antioxidants: Triumphs and pitfalls of preclinical findings and their translational limitations. Chemico-Biological Interactions, 183, 333–340.

USEPA, 1980. United States Environmental Protection Agency, Ambient water quality criteria for chlorinated phenols. Washington, D.C., USA.

USEPA, 2002a. United States Environmental Protection Agency, Toxicological Review of Phenol. Washington D.C, USA.

USEPA, 2002b. United States Environmental Protection Agency, Short-term methods for estimating the chronic toxicity of effluents and receiving waters to fresh water organisms (EPA-821-R-02-013), 4th ed.,Washington DC, USA.

USEPA, 2005. United States Environmental Protection Agency, Guidelines for Carcinogen Risk Assessment, Washington, DC.

USEPA, 2011. United States Environmental Protection Agency, ECOlogical Structure-Activity Relationship Model (ECOSAR), version 1.1., Cincinnati, OH.

USEPA, 2012. United States Environmental Protection Agency, ECOTOXicology database, available online at: http://cfpub.epa.gov/ecotox/ (Accessed, February 2013).

USEPA, 2011. United States Environmental Protection Agency, The Estimation Programs Interface (EPI) Suite, Version 4.1., Cincinnati, OH.

Varagnal, J., 1981. Hydroquinone, resorcinol, and catechol, third ed. In: Grayson,M., Eckroth, D., Bushey, G.J., Eastman, C.I., Klingsberg, A., Spiro, L., Wainwright, M. (Eds.), Kirk-Othmer Encyclopedia of Chemical Technology, vol. 13 JohnWiley & Sons, New York, pp. 39–69.

Ventura, S.P.M., Gonçalves, A.M.M., Gonçalves, F., Coutinho, J.A.P., 2010. Assessing the toxicity on [C3mim][Tf2N] to aquatic organisms of different trophic levels. Aquatic Toxicology, 96, 290-297.

Verhaar, H.J.M., van Leeuwen, C.J., Hermens, J.L.M., 1992. Classifying environmental pollutants. 1: Structure-activity relationships for prediction of aquatic toxicity. Chemosphere, 25, 471–491.

Vermuza, K., Filzmoser, P., 2009. Introduction to multivariate statistical analysis in chemometrics. CRC Press, Taylor&Francis Group, USA.

Vighi, M., Migliorati, S., Monti, G.S., 2009. Toxicity on the luminescent bacterium *Vibrio fischeri* (Beijerinck). I: QSAR equation for narcotics and polar narcotics. Ecotoxicology and Environmental Safety, 72, 154–161.

Vink, S.R., Mikkers, J., Bouwman, T., Marquart, H., Kroese, E.D., 2010. Use of read-across and tiered exposure assessment in risk assessment under REACH – A case study on a phase-in substance. Regulatory Toxicology and Pharmacology, 58, 64–71.

Yao, X.J., Panaye, J.P. Doucet, H.F. Chen, R.S. Zhang, B.T. Fan, Liu, M.C., Hu, Z.D., 2005. Comparative classification study of toxicity mechanisms using support vector machines and radial basis function neural networks. Analytica Chimica Acta, 535, 259-273.

WHO, 1987. World Health Organization, Environmental health criteria 71: pentachlorophenol. Geneva.

WHO, 1989. World Health Organization, Environmental health criteria 93: chlorophenols other than pentachlorophenol. Geneva.

WHO, 1994. World Health Organization, Environmental health criteria 161: phenol. Geneva.

WHO, 1996. World Health Organization, Environmental health criteria 157: hydroquinone. Geneva.

WHO, 2006. World Health Organization, Concise international chemical assessment document 71: resorcinol. Geneva.

Wiegman, S., van Vlaardingen, P.L.A., Bleeker, E.A.J., de Voogt, P., Kraak, M.H.S., 2001. Phototoxicity of azaarene isomers to the marine flagellate *Dunaliella tertiolecta*. Environmental Toxicology and Chemistry, 20, 1544-1550.

Worgan, A.D.P., Dearden, J.C., Edwards, R., Netzeva, T.I., Cronin, M.T.D., 2003. Evaluation of a novel short-term algal toxicity assay by the development of QSARs and inter-species relationships for narcotic chemicals. QSAR and Combinatorial Science, 22, 204-209.

Worth, A.P., Cronin, M.T.D., 2004. Report of the workshop on the validation of QSARs and other computational prediction models. ATLA: Alternatives to Laboratory Animals, 32, 703–706.

Worth, A.P., 2010. The role of QSAR methodology in the regulatory assessment of chemicals. In: Recent Advances in QSAR studies. Puzyn, T., Leszczynski, J., Cronin, M.T.D. (Eds), Springer, New York.

Zhang, X.J., Qin, H.W., Su, L.M., Qin, W.C., Zou, M.Y., Sheng, L.X., Zhao, Y.H., Abraham, M.H., 2010. Interspecies correlations of toxicity to eight aquatic organisms: theoretical considerations. Science of the Total Environment, 408, 4549-4555.

Zhu, H., Tropsha, A., Fourches, D., Varnek, A., Papa, E., Gramatica, P., Oberg, T., Dao, P., Cherkasov, A., Tetko, I.V., 2008. Combinatorial QSAR modeling of chemical toxicants tested against *Tetrahymena pyriformis*. Journal of Chemical Information and Modeling, 48, 766-784.

Zou, J., Han, Y., So, S.S., 2008. Overview of artificial neural networks. In: Artificial Neural Networks, Methods and Applications, Livingstone, D. J. (ed.), Humana press, Totowa, NJ, pp.15-23.

Zupan, J., Gasteiger, J., 1999. Neural networks in chemistry and drug design, Wiley-VCH, Weinheim.

Zvinavashe, E., Murk, A.J., Rietjens, I.M.C.M., 2008. Promises and pitfalls of quantitative structure-activity relationship approaches for predicting metabolism and toxicity. Chemical Research in Toxicology, 21, 2229-2236.

APPENDIX A: PLOTS OF ABSORBANCE VERSUS ALGAL CELLS

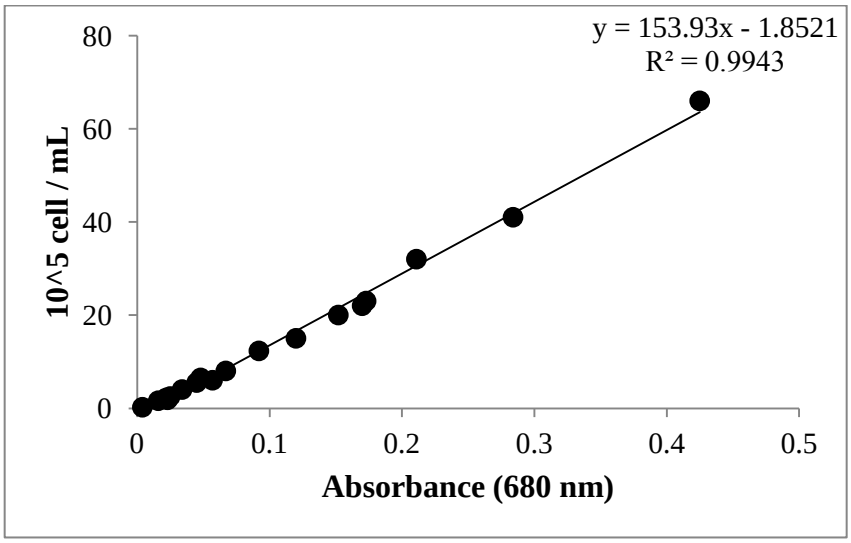


Figure A1. Absorbance versus number of algal cells for *Chlorella vulgaris*

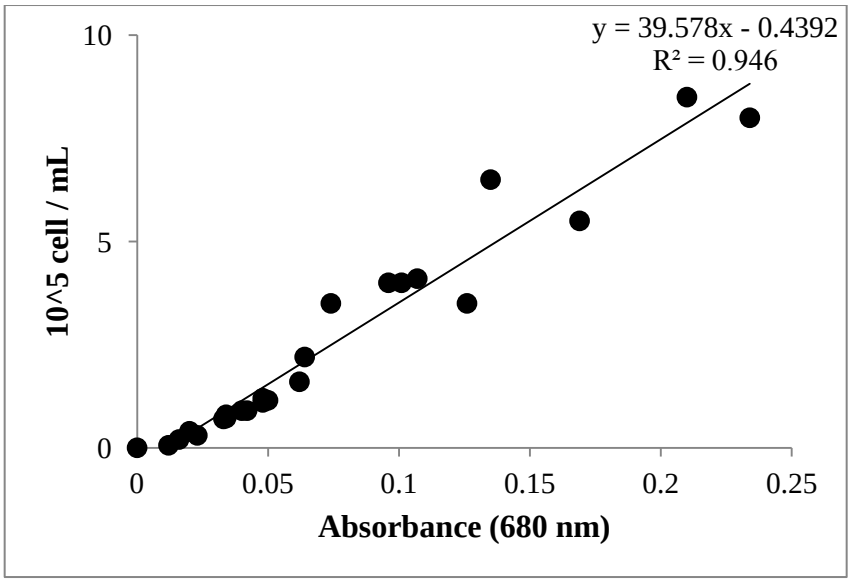


Figure A2. Absorbance versus number of algal cells for *Dunaliella tertiolecta*

APPENDIX B: SOFTWARE USED IN THE THESIS

Table B1. Software used in the thesis and their main are of use

Software	Area of use	Reference
Dragon	Descriptor generation	Ver. 5.4, 2006, Talete, Milano, Italy
CODESSA	Descriptor generation, statistical package	Ver. 2.2, 1996, University of Florida, Gainesville, USA
CODESSA Pro	Descriptor generation, statistical package	Ver. Pro, 2002, University of Florida, Gainesville, USA
CP ANN Modules	CP ANN model construction	© Institute of Chemistry, 1999, Slovenia
Molegro Data Modeller	Statistical package	Ver. 2.6.0., 2011, Molegro ApS, Aarhus, Denmark
ECOSAR	Aquatic toxicity prediction	Ver. 1.1, 2012, United States Environmental Protection Agency
EPISuite	Data retrieval on chemical properties	Ver. 3.20, 2007, United States Environmental Protection Agency
Scientific Workplace	Mathematical software	Ver. 3.0, 1998, MacKichan software, WA, USA
Spartan	Molecule drawing, geometry optimization	Ver. 06, 2006, Wavefunction, Inc., CA, USA
SPSS	Statistical package	Ver. 20.0.0, 2011, SPSS, Inc., Chicago, IL, USA
ToxCalc	Toxicity calculation, hypothesis testing	Ver. 5.0.32, 2009, Tidepool Scientific Software, CA, USA

APPENDIX C: CALIBRATION CURVE AND GC CHROMATOGRAM FOR 2-CHLOROPHENOL

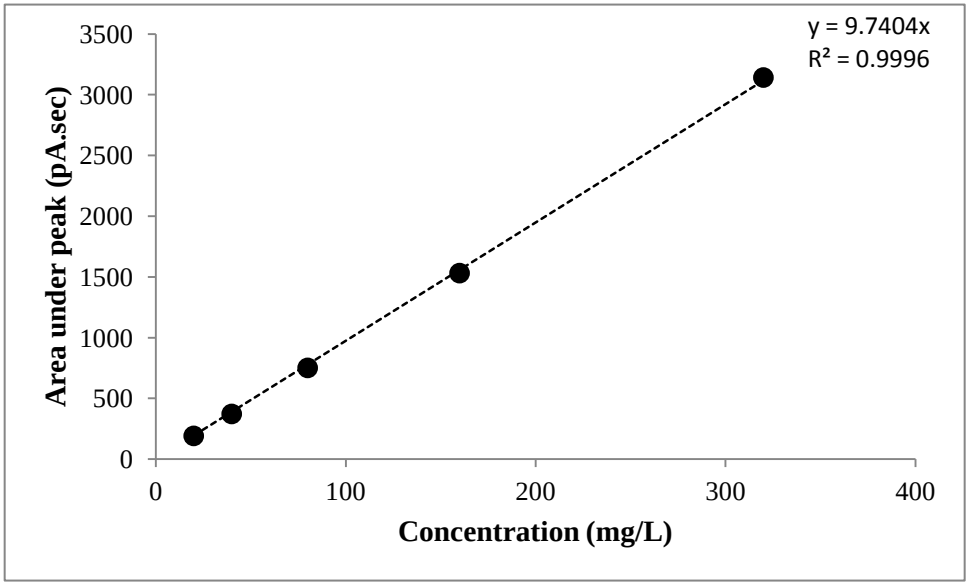


Figure C1. 2-chlorophenol calibration curve for gas chromatographic analysis

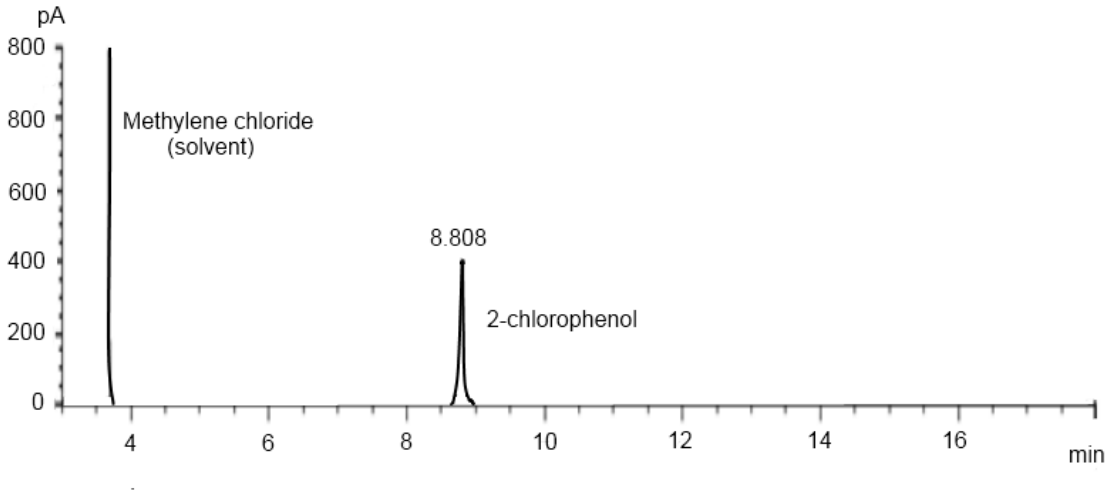


Figure C2. GC chromatogram for 100 mg/L 2-chlorophenol

APPENDIX D: EXTERNAL SET DETAILS OF THE ALGAL QSTRS

Table D1. Descriptor values for the external set compounds for QSARs developed for *D. tertiolecta*

ID	Compound	Log *P*	Mor24m	Mor24p	Mor18u	C2	HATS7e	K1maxa
V01	2-methylphenol	1.95	0.09	0.06	-0.69	-0.14	0.00	9.04
V02	3-methylphenol	1.96	0.07	0.02	-0.59	-0.15	0.00	9.09
V03	4-methylphenol	1.94	0.07	0.03	-0.61	-0.12	0.25	9.08
V04	2,3-dimethylphenol	2.48	0.10	0.06	-1.13	-0.14	0.00	9.03
V05	2,4-dimethylphenol	2.30	0.10	0.05	-0.79	-0.11	0.11	9.04
V06	2,5-dimethylphenol	2.33	0.09	0.03	-0.71	-0.14	0.69	9.05
V07	2,6-dimethylphenol	2.36	0.12	0.06	-0.82	-0.14	0.00	9.03
V08	3,4-dimethylphenol	2.23	0.07	0.02	-0.87	-0.11	0.15	9.10
V09	3,5-dimethylphenol	2.35	0.07	-0.01	-0.62	-0.15	0.00	9.09
V10	2,3,5-trimethylphenol	3.15	0.11	0.04	-1.11	-0.15	0.40	9.04
V11	2,3,6-trimethylphenol	2.67	0.14	0.08	-1.22	-0.14	0.43	9.01
V12	2,4,6-trimethylphenol	2.73	0.13	0.05	-0.89	-0.11	0.07	9.04
V13	2-ethylphenol	2.47	0.12	0.11	-1.08	-0.14	0.14	9.00
V14	3-ethylphenol	2.40	0.09	0.07	-0.65	-0.14	0.29	9.09
V15	4-ethylphenol	2.58	0.10	0.08	-0.62	-0.12	0.26	9.09
V16	Aniline	0.90	0.08	0.07	-0.81	-0.14	0.00	8.71
V17	2-chloroaniline	1.90	0.09	0.09	-0.75	-0.13	0.00	8.73
V18	3-chloroaniline	1.88	0.07	0.08	-0.78	-0.14	0.00	8.70
V19	4-chloroaniline	1.83	0.09	0.09	-0.73	-0.16	0.00	8.67
V20	2,3-dichloroaniline	2.82	0.14	0.09	-0.67	-0.13	0.00	8.71
V21	2,4-dichloroaniline	2.78	0.13	0.12	-0.66	-0.15	0.00	8.72
V22	2,5-dichloroaniline	2.75	0.08	0.10	-0.72	-0.13	0.00	8.72
V23	2,6-dichloroaniline	2.76	0.03	0.11	-0.71	-0.12	0.00	8.76
V24	3,4-dichloroaniline	2.69	0.16	0.10	-0.65	-0.17	0.00	8.67
V25	3,5-dichloroaniline	2.90	0.08	0.10	-0.73	-0.14	0.00	8.68
V26	2,3,4-trichloroaniline	3.33	0.27	0.12	-0.54	-0.16	0.00	8.68
V27	2,4,5-trichloroaniline	3.45	0.21	0.12	-0.60	-0.16	0.00	8.70
V28	2,4,6-trichloroaniline	3.52	0.10	0.14	-0.62	-0.14	0.00	8.72
V29	3,4,5-trichloroaniline	3.32	0.27	0.12	-0.57	-0.18	0.00	8.65
V30	2-methylaniline	1.32	0.08	0.07	-0.91	-0.13	0.00	8.70
V31	3-methylaniline	1.40	0.09	0.05	-0.93	-0.14	0.00	8.79
V32	4-methylaniline	1.39	0.08	0.04	-0.94	-0.11	0.45	8.77
V33	2,3-dimethylaniline	2.17	0.11	0.07	-1.06	-0.14	0.00	8.71
V34	2,4-dimethylaniline	1.68	0.08	0.04	-1.01	-0.10	0.24	8.78
V35	2,5-dimethylaniline	1.83	0.09	0.05	-1.00	-0.13	0.65	8.81
V36	2,6-dimethylaniline	1.84	0.08	0.07	-1.18	-0.13	0.00	8.68
V37	3,4-dimethylaniline	1.84	0.08	0.03	-1.16	-0.11	0.26	8.77
V38	3,5-dimethylaniline	2.17	0.11	0.04	-1.03	-0.14	0.00	8.79
V39	2,4,6-trimethylaniline	2.72	0.09	0.05	-1.25	-0.10	0.14	8.80
V40	2-ethylaniline	1.74	0.08	0.12	-1.17	-0.13	0.13	8.67
V41	3-ethylaniline	2.11	0.11	0.10	-0.97	-0.13	0.36	8.75
V42	4-ethylaniline	1.96	0.10	0.09	-0.98	-0.11	0.30	8.73
V43	2,6-diethylaniline	3.15	0.10	0.15	-1.62	-0.13	0.48	8.66
V44	Phenol	1.46	0.07	0.05	-0.50	-0.14	0.00	9.08
V45	2-chlorophenol	2.15	0.03	0.04	-0.43	-0.13	0.00	9.14
V46	3-chlorophenol	2.50	0.12	0.07	-0.46	-0.15	0.00	9.07

Table D1. Continued

ID	Compound	Log *P*	Mor24m	Mor24p	Mor18u	C2	HATS7e	K1maxa
V47	4-chlorophenol	2.39	0.11	0.06	-0.47	-0.17	0.00	9.05
V48	2,3-dichlorophenol	2.84	0.17	0.07	-0.37	-0.14	0.00	9.12
V49	2,4-dichlorophenol	3.06	0.10	0.07	-0.40	-0.16	0.00	9.10
V50	2,5-dichlorophenol	3.06	0.04	0.07	-0.41	-0.13	0.00	9.13
V51	2,6-dichlorophenol	2.75	0.00	0.05	-0.44	-0.12	0.00	9.17
V52	3,4-dichlorophenol	3.33	0.25	0.09	-0.40	-0.18	0.00	9.04
V53	3,5-dichlorophenol	3.62	0.16	0.11	-0.42	-0.15	0.00	9.06
V54	2,3,4-trichlorophenol	3.80	0.33	0.10	-0.31	-0.17	0.00	9.08
V55	2,3,5-trichlorophenol	3.84	0.21	0.10	-0.35	-0.14	0.00	9.11
V56	2,3,6-trichlorophenol	3.77	0.14	0.07	-0.39	-0.12	0.00	9.15
V57	2,4,5-trichlorophenol	3.72	0.20	0.09	-0.35	-0.17	0.00	9.09
V58	2,4,6-trichlorophenol	3.69	0.10	0.08	-0.42	-0.15	0.00	9.12

Table D2. Toxicity data, training, test and external validation set predictions, leverages and standardized residuals for *D. tertiolecta* models

Compound	pT (mmol/L)	Predictions			Leverages			Standardized residuals		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Training set										
Phenol	-0.24	0.57	0.07	0.26	0.03	0.09	0.15	2.04	0.80	1.17
2-chlorophenol	0.44	0.43	-0.02	0.58	0.06	0.11	0.07	-0.03	-1.19	0.33
4-chlorophenol	0.81	0.68	0.54	0.76	0.02	0.04	0.19	-0.32	-0.70	-0.11
2,3-dichlorophenol	0.98	1.26	0.82	1.41	0.00	0.01	0.03	0.71	-0.42	1.00
2,4-dichlorophenol	1.23	0.72	0.72	1.16	0.02	0.02	0.00	-1.26	-1.30	-0.16
2,6-dichlorophenol	0.98	0.49	0.69	0.89	0.05	0.02	0.03	-1.24	-0.74	-0.22
2,6-dichlorophenol	0.43	0.33	0.23	0.26	0.08	0.07	0.20	-0.26	-0.52	-0.40
3,5-dichlorophenol	1.71	1.07	2.00	1.30	0.00	0.05	0.08	-1.59	0.74	-0.95
2,3,4-trichlorophenol	1.88	2.05	1.59	2.29	0.06	0.01	0.10	0.41	-0.73	0.96
2,3,5-trichlorophenol	1.78	1.48	1.84	1.71	0.00	0.03	0.03	-0.75	0.16	-0.17
2,3,6-trichlorophenol	0.80	1.17	0.94	0.97	0.00	0.01	0.10	0.92	0.36	0.40
2,4,6-trichlorophenol	0.75	0.75	1.16	0.80	0.02	0.00	0.02	0.00	1.05	0.11
3,4,5-trichlorophenol	2.22	2.18	2.47	2.40	0.10	0.11	0.26	-0.11	0.63	0.42
2,3,4,5-tetrachlorophenol	2.59	2.71	2.56	2.69	0.18	0.13	0.20	0.30	-0.08	0.23
2,3,4,6-tetrachlorophenol	2.29	2.03	1.87	1.62	0.05	0.03	0.03	-0.66	-1.07	-1.56
2,3,5,6-tetrachlorophenol	2.32	1.81	1.81	1.57	0.02	0.03	0.07	-1.26	-1.30	-1.74
Pentachlorophenol	3.03	3.17	2.84	2.52	0.29	0.19	0.16	0.35	-0.49	-1.18
1,2,3-trihydroxybenzene	0.59	0.70	0.85	0.33	0.04	0.01	0.07	0.27	0.66	-0.60
Hydroquinone	2.56	2.03	2.03	2.51	0.31	0.54	0.37	-1.32	-1.36	-0.12
Tetrachlorohydroquinone	1.68	2.27	2.29	1.99	0.41	0.39	0.41	1.48	1.57	0.72
4-chlorocatechol	0.86	1.03	1.07	0.68	0.01	0.01	0.08	0.42	0.53	-0.43
3,5-dichlorocatechol	1.19	1.51	1.50	0.45	0.01	0.01	0.10	0.80	0.80	-1.73
Resorcinol	-0.42	0.18	0.38	0.40	0.12	0.05	0.09	1.49	2.06	1.91
4-chlororesorcinol	0.51	0.37	0.63	0.13	0.12	0.03	0.14	-0.35	0.31	-0.87
Test set										
3-chlorophenol	0.91	0.90	0.94	0.78	0.01	0.01	0.08	-0.01	0.08	-0.29
3,4-dichlorophenol	1.43	1.51	1.38	1.58	0.02	0.01	0.13	0.20	-0.14	0.36
2,4,5-trichlorophenol	1.99	1.25	1.44	1.76	0.01	0.01	0.03	-1.85	-1.41	-0.53
Chlorohydroquinone	2.59	2.44	2.39	2.18	0.24	0.29	0.24	-0.39	-0.50	-0.94
Catechol	0.73	1.07	0.41	0.69	0.02	0.04	0.06	0.84	-0.81	-0.10
4,6-dichlororesorcinol	0.16	0.23	0.54	-0.31	0.12	0.03	0.17	0.17	0.97	-1.09

Table D2. Continued

Compound	pEC_{50} (mmol/L)	Predictions			Leverages			Standardized residuals		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
External validation set										
2-methylphenol		-0.07	0.72	0.38	-1.49	0.02	0.05	0.51	1.97	1.16	-3.29
3-methylphenol		-0.13	0.59	-0.71	-0.88	0.03	0.19	0.30	1.81	-1.48	-1.74
4-methylphenol		0.27	0.79	10.98	-0.91	0.02	0.92	0.33	1.30	27.46	-2.75
2,3-dimethylphenol		0.40	0.80	0.41	-6.33	0.01	0.04	0.83	1.01	0.03	-15.65
2,4-dimethylphenol		0.80	0.98	5.03	-2.46	0.01	0.77	0.62	0.46	10.85	-7.58
2,5-dimethylphenol		0.58	0.71	30.23	-1.78	0.02	0.95	0.54	0.33	76.04	-5.48
2,6-dimethylphenol		0.47	0.92	0.54	-2.84	0.01	0.03	0.66	1.12	0.17	-7.70
3,4-dimethylphenol		0.58	0.79	6.09	-3.87	0.03	0.85	0.68	0.53	14.12	-10.32
3,5-dimethylphenol		0.65	0.59	-1.55	-1.22	0.03	0.33	0.35	-0.15	-5.63	-4.34
2,3,5-trimethylphenol		1.00	0.84	17.58	-6.23	0.01	0.94	0.82	-0.41	42.52	-16.81
2,3,6-trimethylphenol		0.98	1.05	19.92	-7.20	0.00	0.95	0.86	0.17	48.57	-19.02
2,4,6-trimethylphenol		1.15	1.21	3.33	-3.54	0.01	0.58	0.71	0.14	5.60	-10.90
2-ethylphenol		0.59	0.91	8.45	-5.51	0.01	0.85	0.82	0.81	20.15	-14.18
3-ethylphenol		0.48	0.74	13.36	-1.51	0.02	0.93	0.41	0.64	33.03	-4.62
4-ethylphenol		0.75	0.91	12.76	-1.00	0.02	0.92	0.33	0.39	30.79	-4.08
Aniline		0.24	0.67	0.79	-0.03	0.03	0.01	0.86	1.06	1.40	-0.62
2-chloroaniline		0.51	0.79	1.44	0.60	0.02	0.01	0.84	0.71	2.38	0.20
3-chloroaniline		0.68	0.58	1.19	0.52	0.03	0.00	0.86	-0.24	1.31	-0.37
4-chloroaniline		1.56	0.61	1.44	1.12	0.03	0.01	0.86	-2.37	-0.31	-1.03
2,3-dichloroaniline		1.38	1.08	1.50	1.57	0.00	0.01	0.84	-0.74	0.31	0.44
2,4-dichloroaniline		1.61	0.91	2.31	1.61	0.01	0.08	0.83	-1.75	1.79	-0.00
2,5-dichloroaniline		0.99	0.75	1.72	0.93	0.02	0.02	0.84	-0.59	1.87	-0.14
2,6-dichloroaniline		0.84	0.51	1.91	0.76	0.05	0.03	0.82	-0.83	2.73	-0.19
3,4-dichloroaniline		1.81	1.00	1.75	2.04	0.01	0.02	0.85	-2.02	-0.15	0.52
3,5-dichloroaniline		1.57	0.67	1.84	1.14	0.02	0.03	0.86	-2.25	0.70	-1.00
2,3,4-trichloroaniline		1.74	1.67	2.34	3.14	0.02	0.08	0.84	-0.17	1.54	3.26
2,4,5-trichloroaniline		1.80	1.32	2.50	2.41	0.01	0.10	0.84	-1.20	1.79	1.41
2,4,6-trichloroaniline		1.60	0.80	3.03	1.99	0.01	0.18	0.82	-2.00	3.65	0.90
3,4,5-trichloroaniline		2.14	1.58	2.40	2.99	0.03	0.09	0.86	-1.40	0.67	1.98
2-methylaniline		-0.01	0.70	0.72	-1.04	0.02	0.02	0.87	1.78	1.88	-2.40
3-methylaniline		0.60	0.74	0.23	-2.03	0.02	0.06	0.85	0.34	-0.96	-6.10
4-methylaniline		0.40	0.84	19.86	-1.79	0.03	0.95	0.85	1.11	49.91	-5.10
2,3-dimethylaniline		0.59	0.86	0.88	-2.72	0.01	0.01	0.88	0.68	0.74	-7.69

Table D2. Continued

Compound	pEC_{50} (mmol/L)	Predictions			Leverages			Standardized residuals		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
2,4-dimethylaniline	0.49	0.92	10.58	-2.73	0.02	0.92	0.86	1.08	25.88	-7.49
2,5-dimethylaniline	0.23	0.78	29.07	-2.95	0.02	0.95	0.85	1.38	73.94	-7.39
2,6-dimethylaniline	0.05	0.74	0.85	-3.78	0.02	0.01	0.90	1.74	2.05	-8.91
3,4-dimethylaniline	1.22	0.88	11.15	-4.24	0.02	0.92	0.88	-0.86	25.45	-12.70
3,5-dimethylaniline	0.64	0.86	-0.08	-3.18	0.01	0.10	0.86	0.55	-1.86	-8.88
2,4,6-trimethylaniline	0.82	1.01	6.32	-5.54	0.02	0.83	0.89	0.47	14.10	-14.79
2-ethylaniline	0.39	0.69	8.04	-3.67	0.02	0.83	0.90	0.75	19.61	-9.45
3-ethylaniline	0.93	0.89	17.48	-2.14	0.01	0.94	0.86	-0.09	42.44	-7.14
4-ethylaniline	1.14	0.96	14.61	-1.94	0.02	0.93	0.87	-0.45	34.54	-7.16
2,6-diethylaniline	0.56	0.86	24.61	-8.67	0.01	0.95	0.92	0.74	61.66	-21.47
Phenol	-0.32	0.57	0.07	0.26	0.03	0.08	0.17	2.24	1.00	1.35
2-chlorophenol	0.39	0.43	-0.02	0.58	0.05	0.09	0.19	0.08	-1.07	0.44
3-chlorophenol	1.05	0.90	0.94	0.78	0.01	0.01	0.14	-0.36	-0.27	-0.61
4-chlorophenol	0.61	0.68	0.54	0.76	0.02	0.03	0.18	0.18	-0.19	0.35
2,3-dichlorophenol	1.17	1.26	0.82	1.41	0.00	0.01	0.17	0.22	-0.92	0.55
2,4-dichlorophenol	1.30	0.72	0.72	1.16	0.02	0.02	0.13	-1.45	-1.48	-0.33
2,5-dichlorophenol	1.65	0.49	0.69	0.89	0.04	0.02	0.17	-2.90	-2.45	-1.77
2,6-dichlorophenol	1.01	0.33	0.23	0.26	0.07	0.06	0.27	-1.69	-2.00	-1.74
3,4-dichlorophenol	1.87	1.51	1.38	1.58	0.02	0.01	0.18	-0.91	-1.27	-0.67
3,5-dichlorophenol	1.89	1.07	2.00	1.30	0.00	0.04	0.14	-2.04	0.28	-1.36
2,3,4-trichlorophenol	1.68	2.05	1.59	2.29	0.05	0.01	0.21	0.92	-0.21	1.43
2,3,5-trichlorophenol	1.94	1.48	1.84	1.71	0.00	0.03	0.17	-1.15	-0.25	-0.54
2,3,6-trichlorophenol	1.39	1.17	0.94	0.97	0.00	0.01	0.22	-0.56	-1.15	-0.97
2,4,5-trichlorophenol	1.42	1.25	1.44	1.76	0.01	0.01	0.16	-0.42	0.06	0.80
2,4,6-trichlorophenol	1.54	0.75	1.16	0.80	0.02	0.00	0.15	-1.99	-0.99	-1.73
Hat value (h*)					**0.37**	**0.37**	**0.50**	**Standard residual limit: ±2.50**		

Table D3. Descriptor values for the external set compounds for QSTRs developed for *C. vulgaris*

ID	Compound	Log *P*	*p*Ka	Log *D*	$E_{(HOMO-LUMO)}$ (eV)	Tm	nHBonds
V01	2-methylphenol	1.95	10.30	1.95	9.34	3.62	0
V02	3-methylphenol	1.96	10.10	1.96	9.38	4.02	0
V03	4-methylphenol	1.94	10.30	1.94	9.28	4.21	0
V04	2,3-dimethylphenol	2.48	10.50	2.48	9.28	4.33	0
V05	2,4-dimethylphenol	2.30	10.60	2.30	9.17	4.62	0
V06	2,5-dimethylphenol	2.33	10.40	2.33	9.24	4.53	0
V07	2,6-dimethylphenol	2.36	10.60	2.36	9.26	4.13	0
V08	3,4-dimethylphenol	2.23	10.40	2.23	9.22	4.78	0
V09	3,5-dimethylphenol	2.35	10.20	2.35	9.34	4.75	0
V10	2,3,5-trimethylphenol	3.15	10.70	3.15	9.20	5.10	0
V11	2,3,6-trimethylphenol	2.67	10.80	2.67	9.16	4.86	0
V12	2,4,6-trimethylphenol	2.73	10.90	2.73	9.11	5.03	0
V13	2-ethylphenol	2.47	10.20	2.47	9.33	4.65	0
V14	3-ethylphenol	2.40	9.90	2.39	9.40	5.07	0
V15	4-ethylphenol	2.58	10.00	2.58	9.33	5.38	0
V16	Aniline	0.90	4.60	0.90	9.03	3.06	0
V17	2-chloroaniline	1.90	2.66	1.90	8.79	4.43	0
V18	3-chloroaniline	1.88	3.52	1.88	8.88	5.19	0
V19	4-chloroaniline	1.83	3.98	1.83	8.69	5.59	0
V20	2,3-dichloroaniline	2.82	1.76	2.82	8.69	5.77	0
V21	2,4-dichloroaniline	2.78	2.00	2.78	8.53	6.18	0
V22	2,5-dichloroaniline	2.75	2.05	2.75	8.61	5.92	0
V23	2,6-dichloroaniline	2.76	0.42	2.76	8.63	5.28	0
V24	3,4-dichloroaniline	2.69	2.97	2.69	8.56	6.67	0
V25	3,5-dichloroaniline	2.90	2.51	2.90	8.79	6.46	0
V26	2,3,4-trichloroaniline	3.33	1.58	3.33	8.45	6.93	0
V27	2,4,5-trichloroaniline	3.45	1.09	3.45	8.35	7.05	0
V28	2,4,6-trichloroaniline	3.52	-0.03	3.52	8.42	6.62	0
V29	3,4,5-trichloroaniline	3.32	2.21	3.32	8.49	7.41	0
V30	2-methylaniline	1.32	4.44	1.32	8.95	3.65	0
V31	3-methylaniline	1.40	4.69	1.40	8.98	3.95	0
V32	4-methylaniline	1.39	5.10	1.39	8.86	4.11	0
V33	2,3-dimethylaniline	2.17	4.70	2.17	8.93	4.28	0
V34	2,4-dimethylaniline	1.68	4.89	1.68	8.81	4.59	0
V35	2,5-dimethylaniline	1.83	4.53	1.83	8.88	4.55	0
V36	2,6-dimethylaniline	1.84	3.95	1.84	8.91	4.22	0
V37	3,4-dimethylaniline	1.84	5.28	1.84	8.84	4.65	0
V38	3,5-dimethylaniline	2.17	4.79	2.17	8.96	4.70	0
V39	2,4,6-trimethylaniline	2.72	4.38	2.72	8.79	5.06	0
V40	2-ethylaniline	1.74	4.30	1.74	8.96	4.45	0
V41	3-ethylaniline	2.11	4.70	2.11	8.98	4.94	0
V42	4-ethylaniline	1.96	5.00	1.96	8.89	5.22	0

Table D3. Continued

ID	Compound	Log *P*	*p*Ka	Log *D*	$E_{(HOMO-LUMO)}$ (eV)	Tm	nHBonds
V43	2,6-diethylaniline	3.15	4.00	3.15	8.91	5.84	0
V44	Phenol	1.46	9.99	1.46	9.47	3.05	0
V45	2-chlorophenol	2.15	8.56	2.04	9.07	4.61	0
V46	3-chlorophenol	2.50	9.12	2.47	9.25	5.02	0
V47	4-chlorophenol	2.39	9.41	2.37	9.06	5.15	0
V48	2,3-dichlorophenol	2.84	7.70	2.36	8.93	5.78	0
V49	2,4-dichlorophenol	3.06	7.89	2.70	8.82	5.99	0
V50	2,5-dichlorophenol	3.06	7.51	2.45	8.77	5.82	0
V51	2,6-dichlorophenol	2.75	6.79	1.51	8.95	5.46	0
V52	3,4-dichlorophenol	3.33	8.63	3.24	8.84	6.20	0
V53	3,5-dichlorophenol	3.62	8.18	3.40	9.15	6.09	0
V54	2,3,4-trichlorophenol	3.80	7.59	3.25	8.68	6.67	0
V55	2,3,5-trichlorophenol	3.84	7.23	3.00	8.65	6.62	0
V56	2,3,6-trichlorophenol	3.77	5.80	1.57	8.80	6.36	0
V57	2,4,5-trichlorophenol	3.72	7.40	3.02	8.56	6.71	0
V58	2,4,6-trichlorophenol	3.69	6.23	1.91	8.68	6.50	0

Table D4. Toxicity data, training, test and external validation set predictions, leverages and standardized residuals for *C. vulgaris* models

Compound	pT (mmol/L)	Predictions			Leverages				Standardized residuals		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3		Model 1	Model 2	Model 3
Training set											
Phenol		-0.60	-0.03	-0.51	-0.69	0.182	0.239	0.297	1.62	0.25	-0.29
2-chlorophenol		0.17	0.65	0.39	0.28	0.026	0.045	0.066	1.37	0.65	0.32
3-chlorophenol		0.36	0.66	0.63	0.53	0.110	0.020	0.034	0.87	0.79	0.53
2,3-dichlorophenol		1.09	1.11	1.07	1.00	0.010	0.000	0.006	0.05	-0.05	-0.27
2,4-dichlorophenol		1.24	1.31	1.20	1.14	0.007	0.001	0.005	0.19	-0.14	-0.34
2,6-dichlorophenol		0.88	0.99	0.89	0.80	0.007	0.004	0.013	0.31	0.02	-0.24
3,4-dichlorophenol		1.47	1.43	1.32	1.27	0.030	0.005	0.008	-0.11	-0.45	-0.64
3,5-dichlorophenol		1.67	1.30	1.25	1.19	0.229	0.003	0.006	-1.06	-1.23	-1.48
2,3,4-trichlorophenol		1.64	1.79	1.59	1.56	0.055	0.024	0.025	0.42	-0.16	-0.27
2,3,5-trichlorophenol		1.86	1.83	1.56	1.52	0.053	0.021	0.023	-0.11	-0.90	-1.06
2,3,6-trichlorophenol		1.51	1.36	1.41	1.37	0.008	0.010	0.012	-0.44	-0.29	-0.45
2,4,5-trichlorophenol		1.67	1.85	1.61	1.58	0.047	0.026	0.028	0.52	-0.18	-0.29
3,4,5-trichlorophenol		2.18	1.87	1.68	1.66	0.092	0.034	0.035	-0.90	-1.46	-1.64
2,3,4,5-tetrachlorophenol		2.34	2.04	1.90	1.89	0.077	0.063	0.065	-0.86	-1.30	-1.41
2,3,5,6-tetrachlorophenol		1.43	1.63	1.76	1.73	0.051	0.042	0.044	0.58	0.96	0.95
Pentachlorophenol		1.45	1.93	2.05	2.05	0.113	0.087	0.090	1.35	1.76	1.86
1,2,3-trihydroxybenzene		1.28	-0.06	0.05	1.14	0.115	0.083	0.634	-3.84	-3.61	-0.46
Chlorohydroquinone		1.13	0.58	0.74	0.64	0.184	0.011	0.023	-1.58	-1.16	-1.52
Tetrachlorohydroquinone		1.46	2.01	1.82	1.80	0.219	0.050	0.052	1.56	1.04	1.05
Catechol		0.27	-0.05	-0.24	0.21	0.158	0.164	0.188	-0.90	-1.48	-0.16
4-chlorocatechol		1.13	0.59	0.86	1.38	0.072	0.005	0.143	-1.56	-0.82	0.77
Resorcinol		-0.49	-0.24	-0.10	-0.25	0.179	0.131	0.170	0.72	1.16	0.77
4-chlororesorcinol		0.27	0.55	0.75	0.66	0.043	0.010	0.022	0.80	1.43	1.23
4,6-dichlororesorcinol		1.02	1.08	1.36	1.31	0.046	0.007	0.010	0.14	0.98	0.89
Test set											
4-chlorophenol		0.46	0.78	0.71	0.61	0.022	0.012	0.023	0.92	0.74	0.49
2,5-dichlorophenol		1.12	1.34	1.10	1.03	0.004	0.000	0.005	0.66	-0.05	-0.27
2,4,6-trichlorophenol		1.53	1.57	1.49	1.45	0.015	0.014	0.016	0.10	-0.12	-0.26
2,3,4,6-tetrachlorophenol		1.45	1.70	1.82	1.81	0.042	0.045	0.046	0.72	1.11	1.12
Hydroquinone		0.02	0.02	-0.02	-0.16	0.216	0.095	0.121	0.01	-0.11	-0.57
3,5-dichlorocatechol		1.95	1.09	1.34	1.90	0.009	0.006	0.157	-2.46	-1.79	-0.15

Table D4. Continued

Compound	pEC_{50} (mmol/L)	Predictions			Leverages			Standardized residuals		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
External validation set										
2-methylphenol		-0.07	0.32	-0.18	-0.33	0.100	0.120	0.149	1.10	-0.32	-0.82
3-methylphenol		-0.13	0.28	0.05	-0.09	0.117	0.084	0.108	1.18	0.54	0.14
4-methylphenol		0.27	0.36	0.16	0.03	0.077	0.068	0.091	0.27	-0.32	-0.75
2,3-dimethylphenol		0.40	0.63	0.23	0.10	0.101	0.059	0.080	0.66	-0.50	-0.93
2,4-dimethylphenol		0.80	0.64	0.40	0.29	0.048	0.039	0.056	-0.46	-1.17	-1.61
2,5-dimethylphenol		0.58	0.59	0.35	0.23	0.075	0.045	0.063	0.03	-0.69	-1.10
2,6-dimethylphenol		0.47	0.59	0.12	-0.02	0.085	0.074	0.098	0.33	-1.04	-1.52
3,4-dimethylphenol		0.58	0.56	0.49	0.38	0.063	0.030	0.045	-0.05	-0.26	-0.62
3,5-dimethylphenol		0.65	0.51	0.48	0.37	0.118	0.031	0.047	-0.39	-0.51	-0.89
2,3,5-trimethylphenol		1.00	1.02	0.68	0.58	0.134	0.014	0.026	0.07	-0.95	-1.31
2,3,6-trimethylphenol		0.98	0.83	0.54	0.43	0.065	0.025	0.040	-0.43	-1.30	-1.71
2,4,6-trimethylphenol		1.15	0.90	0.64	0.54	0.051	0.017	0.029	-0.71	-1.50	-1.91
2-ethylphenol		0.59	0.58	0.41	0.30	0.127	0.037	0.054	-0.04	-0.52	-0.90
3-ethylphenol		0.48	0.48	0.66	0.56	0.150	0.015	0.027	0.01	0.53	0.26
4-ethylphenol		0.75	0.63	0.84	0.75	0.135	0.005	0.014	-0.34	0.26	0.02
Aniline		0.24	0.08	-0.50	-0.68	0.139	0.177	0.211	-0.47	-2.18	-2.88
2-chloroaniline		0.51	0.77	0.29	0.17	0.056	0.051	0.071	0.76	-0.64	-1.07
3-chloroaniline		0.68	0.69	0.73	0.64	0.040	0.011	0.021	0.02	0.15	-0.12
4-chloroaniline		1.56	0.83	0.96	0.89	0.099	0.001	0.008	-2.08	-1.76	-2.11
2,3-dichloroaniline		1.38	1.31	1.07	1.00	0.011	0.000	0.005	-0.19	-0.92	-1.19
2,4-dichloroaniline		1.61	1.44	1.31	1.25	0.058	0.004	0.007	-0.49	-0.89	-1.11
2,5-dichloroaniline		0.99	1.36	1.15	1.09	0.034	0.000	0.005	1.04	0.48	0.31
2,6-dichloroaniline		0.84	1.34	0.78	0.69	0.026	0.008	0.018	1.42	-0.17	-0.45
3,4-dichloroaniline		1.81	1.37	1.59	1.56	0.053	0.021	0.023	-1.27	-0.65	-0.79
3,5-dichloroaniline		1.57	1.27	1.47	1.43	0.001	0.012	0.014	-0.86	-0.30	-0.44
2,3,4-trichloroaniline		1.74	1.78	1.74	1.72	0.060	0.036	0.037	0.11	0.00	-0.07
2,4,5-trichloroaniline		1.80	1.92	1.81	1.79	0.091	0.043	0.045	0.34	0.03	-0.02
2,4,6-trichloroaniline		1.60	1.90	1.56	1.52	0.067	0.019	0.020	0.84	-0.12	-0.24
3,4,5-trichloroaniline		2.14	1.74	2.02	2.02	0.048	0.069	0.072	-1.15	-0.35	-0.39
2-methylaniline		-0.01	0.35	-0.17	-0.32	0.093	0.118	0.147	1.03	-0.46	-0.97
3-methylaniline		0.60	0.36	0.01	-0.13	0.077	0.090	0.115	-0.68	-1.74	-2.29
4-methylaniline		0.40	0.46	0.10	-0.03	0.104	0.076	0.099	0.18	-0.87	-1.35
2,3-dimethylaniline		0.59	0.78	0.20	0.08	0.013	0.062	0.084	0.55	-1.13	-1.61

Table D4. Continued

Compound	pEC_{50} (mmol/L)	Predictions			Leverages			Standardized residuals		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
2,4-dimethylaniline	0.49	0.65	0.38	0.26	0.078	0.041	0.059	0.45	-0.33	-0.71
2,5-dimethylaniline	0.23	0.66	0.36	0.24	0.044	0.043	0.061	1.23	0.38	0.04
2,6-dimethylaniline	0.05	0.64	0.17	0.04	0.039	0.067	0.090	1.69	0.34	-0.04
3,4-dimethylaniline	1.22	0.70	0.42	0.30	0.052	0.037	0.054	-1.47	-2.37	-2.87
3,5-dimethylaniline	0.64	0.75	0.44	0.33	0.013	0.034	0.051	0.33	-0.58	-0.97
2,4,6-trimethylaniline	0.82	1.18	0.66	0.56	0.001	0.016	0.028	1.02	-0.48	-0.82
2-ethylaniline	0.39	0.55	0.30	0.18	0.043	0.050	0.070	0.45	-0.27	-0.66
3-ethylaniline	0.93	0.71	0.59	0.48	0.016	0.021	0.035	-0.63	-1.01	-1.40
4-ethylaniline	1.14	0.71	0.75	0.66	0.030	0.010	0.020	-1.22	-1.15	-1.51
2,6-diethylaniline	0.56	1.29	1.11	1.04	0.025	0.000	0.005	2.07	1.60	1.50
Phenol	-0.32	-0.04	-0.51	-0.69	0.142	0.179	0.213	0.81	-0.57	-1.16
2-chlorophenol	0.39	0.60	0.39	0.28	0.026	0.040	0.057	0.59	-0.01	-0.37
3-chlorophenol	1.05	0.65	0.63	0.53	0.089	0.018	0.030	-1.14	-1.23	-1.62
4-chlorophenol	0.61	0.77	0.71	0.61	0.021	0.012	0.023	0.46	0.28	0.00
2,3-dichlorophenol	1.17	0.88	1.07	1.00	0.006	0.000	0.005	-0.83	-0.30	-0.54
2,4-dichlorophenol	1.30	1.14	1.20	1.14	0.000	0.001	0.005	-0.47	-0.31	-0.52
2,5-dichlorophenol	1.65	1.06	1.10	1.03	0.012	0.000	0.005	-1.68	-1.61	-1.93
2,6-dichlorophenol	1.01	0.45	0.89	0.80	0.070	0.003	0.012	-1.59	-0.35	-0.63
3,4-dichlorophenol	1.87	1.39	1.32	1.27	0.020	0.005	0.007	-1.38	-1.63	-1.89
3,5-dichlorophenol	1.89	1.19	1.25	1.19	0.143	0.002	0.006	-1.99	-1.88	-2.18
2,3,4-trichlorophenol	1.68	1.53	1.59	1.56	0.013	0.021	0.023	-0.42	-0.26	-0.38
2,3,5-trichlorophenol	1.94	1.44	1.56	1.52	0.015	0.019	0.021	-1.43	-1.12	-1.30
2,3,6-trichlorophenol	1.39	0.60	1.41	1.37	0.096	0.009	0.011	-2.24	0.06	-0.08
2,4,5-trichlorophenol	1.42	1.53	1.61	1.58	0.035	0.023	0.025	0.33	0.57	0.51
2,4,6-trichlorophenol	1.54	0.88	1.49	1.45	0.089	0.014	0.016	-1.91	-0.16	-0.29
Hat value (h*)					**0.375**	**0.250**	**0.375**	**Standard residual limit: ±2.50**		