Organic Micropollutants Removal from Water by Oxidation and Other Processes:
QSAR Models, Decision Support System and Hybrids of Processes

Dissertation by

Sairam Sudhakaran

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

King Abdullah University of Science and Technology
Thuwal, Kingdom of Saudi Arabia

(August, 2013)

EXAMINATION COMMITTEE APPROVALS FORM

EXAMINATION COMMITTEE	NAMES
Committee Chairperson:	Prof. Gary Amy
Committee Member:	Prof. Suzana Nunes
Committee Member:	Prof. Niveen Khashab
Committee Member:	Prof. Shane Snyder
Committee Member:	Dr. Sabine Lattemann

ABSTRACT

Organic Micropollutants Removal from Water by Oxidation and Other Processes:

QSAR Models, Decision Support System, and Hybrids of Processes

Sairam Sudhakaran

The presence of organic micropollutants (OMPs) in water is of great environmental concern. OMPs such as endocrine disruptors and certain pharmaceuticals have shown alarming effects on aquatic life. OMPs are included in the priority list of contaminants in several government directorate frameworks. The low levels of OMPs concentration (ng/L to µg/L) force the use of sophisticated analytical instruments. Although, the techniques to detect OMPs are progressing, the focus of current research is only on limited, important OMPs due to the high amount of time, cost and effort involved in analyzing them. Alternatively, quantitative structure activity relationship (QSAR) models help to screen processes and propose appropriate options without considerable experimental effort. QSAR models are well-established in regulatory bodies as a method to screen toxic chemicals.

The goal of the present thesis was to develop QSAR models for OMPs removal by oxidation. Apart from the QSAR models, a decision support system (DSS) based on multi-criteria analysis (MCA) involving socio-economic-technical and sustainability aspects was developed. Also, hybrids of different water treatment processes were studied to propose a sustainable water treatment train for OMPs removal.

In order to build the QSAR models, the ozone/hydroxyl radical rate constants or percent removals of the OMPs were compiled. Several software packages were used to compute the chemical properties of OMPs and perform statistical analyses. For DSS, MCA was used since it allows the comparison of qualitative (non-monetary, non-metric) and quantitative criteria (e.g., costs). Quadrant plots were developed to study the hybrid of natural and advanced water treatment processes.

The QSAR models satisfied both chemical and statistical criteria. The DSS resulted in natural treatment and ozonation as the preferred processes for OMPs removal.

The QSAR models can be used as a screening tool for OMPs removal by oxidation. Moreover, the QSAR - defining molecular descriptors help in detailed understanding of oxidation. The DSS can be considered as an aid in assessing a multi-barrier approach to remove OMPs. Hybrids of natural and advanced treatment processes help to develop a more sustainable multi-barrier approach for OMPs removal.

Key words: Organic Micropollutants, Oxidation, QSAR, DSS, Hybrids

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor Prof. Gary Amy for providing me an opportunity to do my doctoral studies under his tutelage. His constant encouragement, novel ideas and silent inspiration was very helpful to steer my research.

Also, the co-supervision from collaborators during the study, Prof. James Calvin, Dr. Sabine Lattemann, and Prof. Sung Kyu (Andrew) Maeng are deeply appreciated. Prof. James Calvin with his fundamental knowledge of statistics helped me utilize the SPSS software efficiently and understand the statistical concepts related to QSAR modelling. Dr. Sabine Lattemann, with her meticulous thinking and an eye for details was very helpful while developing the decision support system and thesis review. Prof. Sung Kyu (Andrew) Maeng was very supportive in the scientific writing and knowledge related to the hybrid systems work. Special thanks to my senior colleagues from UNESCO-IHE, Dr. Victor Yangali-Quintanilla, Dr. David de Ridder and Mr. Jeong Min Yee for their inputs on QSAR modeling and Decision Support System. The scientific discussions with Prof. Luigi Cavallo, Dr. AbdulRajjak S Shaikh, and Dr. Farhan P. Syed (catalysis center, KAUST) are also acknowledged.

I would also like to thank all my WDRC friends, KAUST IT and administration staff for their support to make our office life quite comfortable and also the skillslab people for their support on scientific writing.

I am grateful to my parents and wife for their constant encouragement and well-timed enquiry regarding my research.

TABLE OF CONTENTS

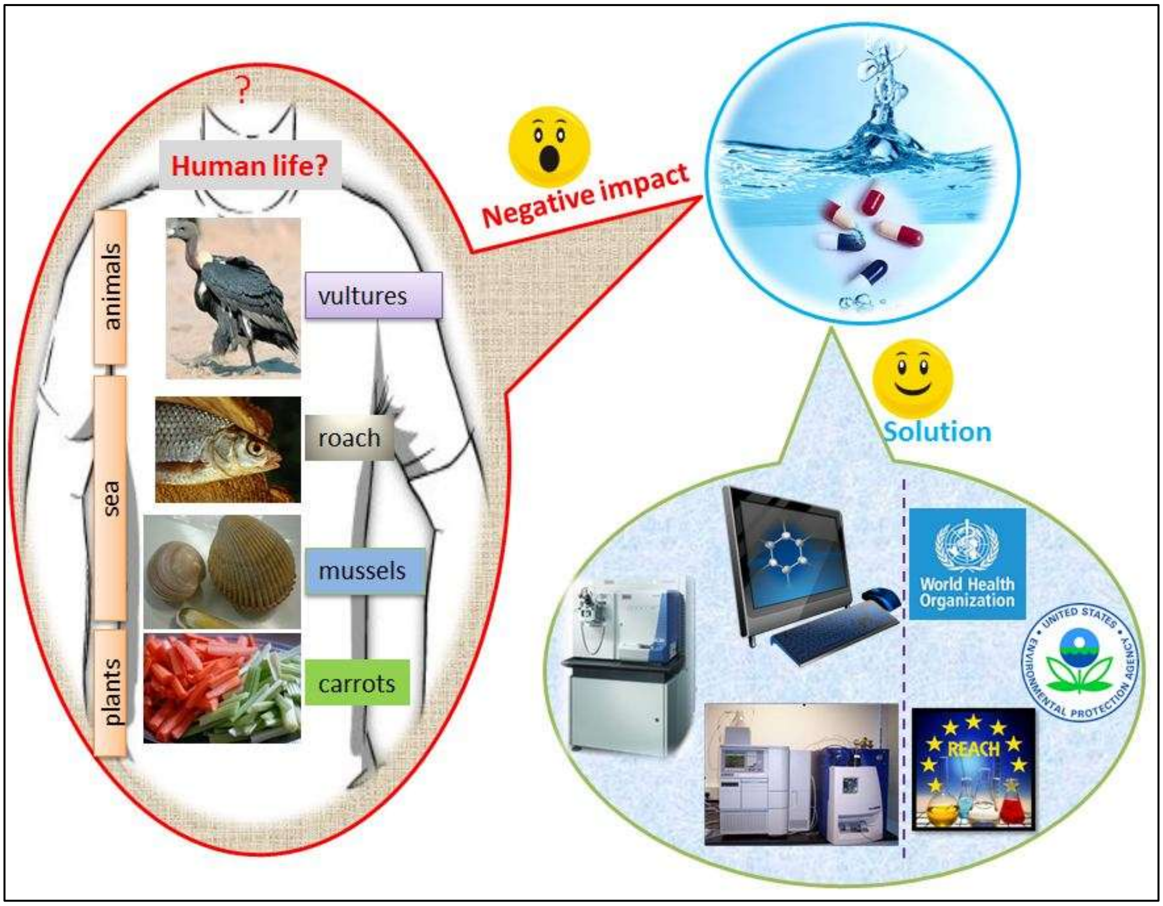
LIST OF ABBREVIATIONS

PPCPs: Pharmaceuticals and Personal Care Products
OMPs: Organic Micropollutants
QSAR: Quantitative Structure Activity Relationship
CRW: Colorado River Water
PRW: Passaic River Water
SRW: Suwannee River Water
ORW: Ohio River Water
MON: Mean Oxidation Number
EA: Electron Affinity
O/C: Oxygen to Carbon Ratio
IP: Ionisation Potential
EA: Electron Affinity
DBE: Double Bond Equivalence
WPSA: Weakly Polar Component of Solvent Accessible Surface Area
k_{OH}, k_{O3}: Hydroxyl Radical Rate Constants, Ozone Rate Constants
AOP: Advanced Oxidation Process
RO: Reverse Osmosis
NF: Nano Filtration
GAC: Granular Activated Carbon
CW: Constructed Wetlands
RBF: River Bank Filtration
ARR: Aquifer Recharge and Recovery
DSS: Decision Support System
MCA: Multi Criteria Analysis
NOM: Natural Organic Matter
DOC: Dissolved Organic Carbon
BPs: By-Products
R/M: Reliability and Maintenance
PM: Potential for Modification

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION



1.1. GLOBAL IMPACT OF ORGANIC MICROPOLLUTANTS

The world pharmaceutical market was estimated at US$ 875 billion in 2010 with an upwelling from US$ 647 billion in 2005 to US$ 875 billion in 2010, corresponding to an increase of 35.2% [1]. Pharmaceuticals and personal care products (PPCPs) are designed to have a physiological effect on humans and animals in trace concentrations. These PPCPs are used yearly with different purposes such as prevention, diagnosis, cure and mitigation of diseases or just to improve the state of health of not only humans but also animals. Although the PPCPs are present at low concentrations (ng/L to µg/L) levels, their persistence against degradation is one of the key properties of PPCPs leading to their classification as pollutants. They retain their chemical structure long enough to do their therapeutic work and because of their continuous input they can remain in the environment for a long time and their presence is therefore considered dangerous in both low and high concentrations [2]. The PPCPs, pesticides, and endocrine disrupting compounds all contain organic moieties in their structure and hence can be broadly classified as organic micropollutants (OMPs). OMPs represent a great challenge as their effects depend on both the level and timing of exposure, especially when exposure occurs during the developmental stage. Also, several studies have shown the adverse effects of OMPs on animal life and these may occur in humans if exposed at a vulnerable time and at concentrations leading to biological modifications [3-5]. Also, exposure to a mixture of OMPs at the same time may result in further adverse effects. There is a global transport of these OMPs through natural processes (e.g., water and air) as well as commercial production, leading to worldwide exposure of human and wildlife to OMPs [5]. Studies have shown the presence of OMPs in the environment more than 30 years ago and the main entry pathways for the OMPs are through bathing, excretion and improper disposal of pharmaceutical waste. To date, around 15 to 25 pharmaceuticals have been detected in treated drinking water worldwide [6, 7]. The amount of OMPs are more in untreated water sources, such as wastewater, surface water and groundwater [8].

The U.S.EPA, in its Unregulated Contaminant Monitoring Rule 3 (2013-2015) signed on April 16, 2012, has included 30 contaminants (28 OMPs and 2 viruses) (http://water.epa.gov/lawsregs/rulesregs/sdwa/ucmr/ucmr3/methods.cfm). There have been several international initiatives to address the impact of OMPs. These include REACH (**R**egistration, **E**valuation, **A**uthorisation and Restriction of **Ch**emical substances) which is the European Community Regulation on chemicals and their safe use which has been in effect since 1st June, 2007, and the Strategic Approach to International Chemicals Management (SAICM) in 2006 [5].

1.2. WATER TREATMENT PROCESSES FOR OMPs REMOVAL

As the regulations with regard to the OMPs in water are getting more stringent, great attention has been focused on OMPs removal by water treatment process experts. Conventional treatment processes which comprise activated sludge, coagulation, settling, filtration (biological, sand) with chlorination can remove about 50% of OMPs whereas the advanced treatment processes such as ozonation, advanced oxidation, activated carbon and membranes can remove more than 99% of OMPs. *Figure 1.1* shows the removal of OMPs in drinking water treatment plants. These results are from bench-scale, pilot-scale and/or full-scale studies [9]. *Figure 1.2* shows the removal of OMPs from conventional and advanced wastewater treatment processes. These include raw, primary settled sewage, activated sludge, secondary treated and treated effluents.

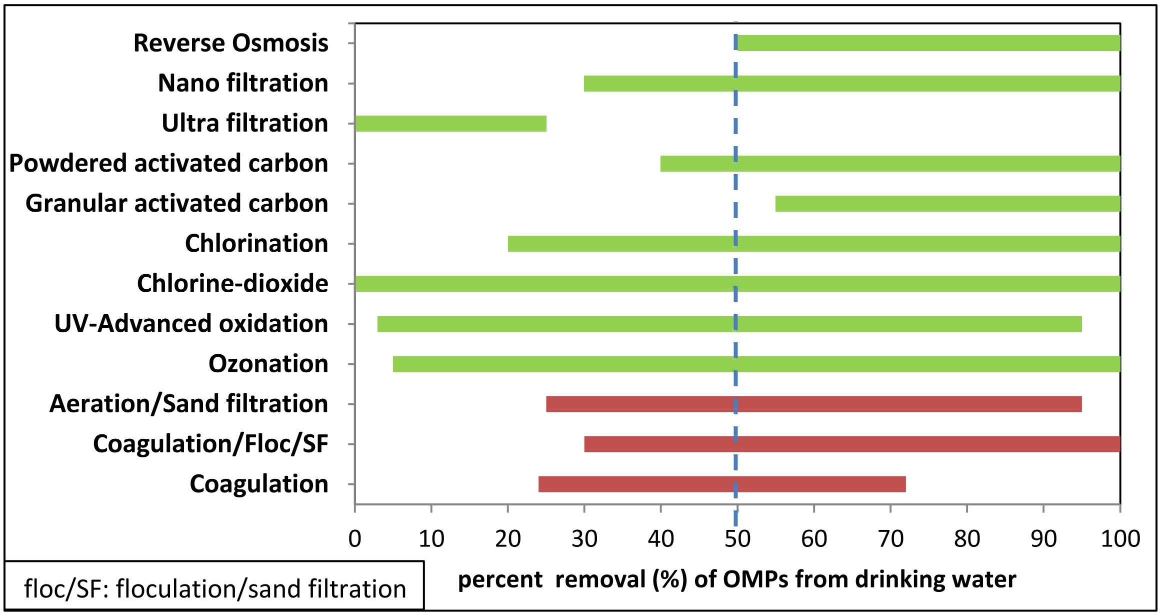


Figure 1.1: OMPs removal from drinking water by conventional and advanced processes [9]

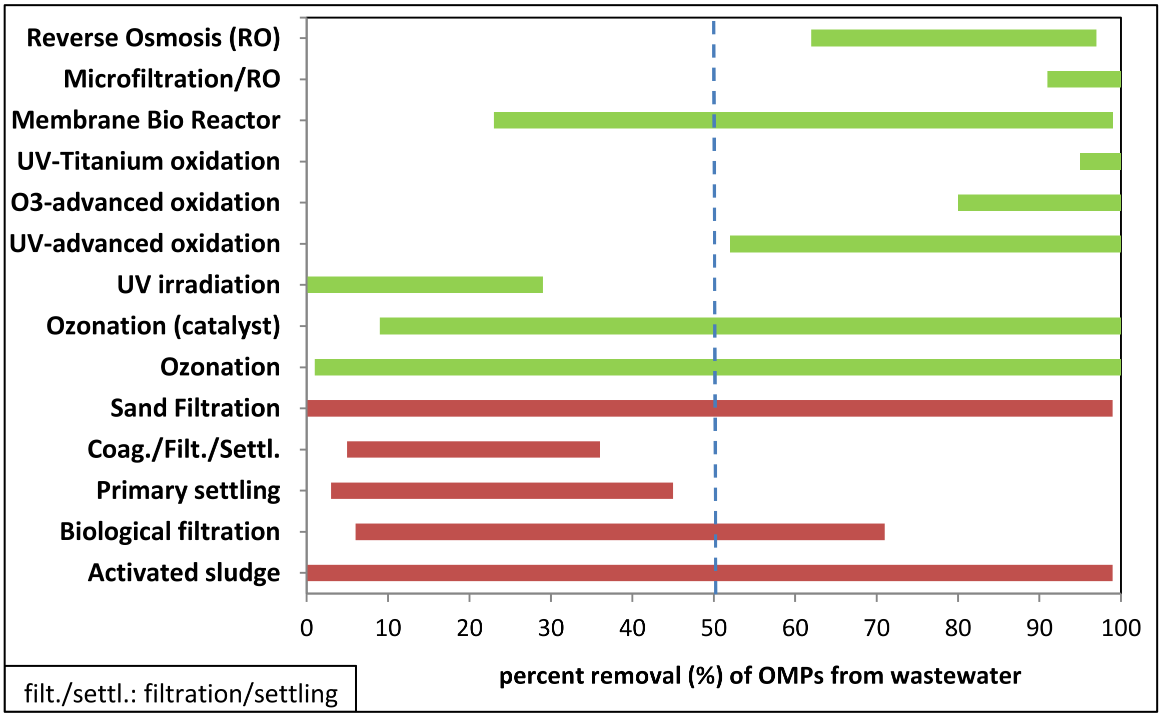


Figure 1.2: OMPs removal from wastewater by conventional and advanced processes [9]

Several countries do not have monitoring programs for OMPs in water due to practical difficulties, such as high costs and unavailability of analytical technologies to detect the wide range of OMPs and their metabolites. The majority of present data on OMPs in drinking/waste water come from targeted research projects, investigations and surveys, which were used to either develop or fine-tune the detection and analytical methods. Hence, a modelling approach such as quantitative structure activity relationship (QSAR) models can be a suitable scientific alternative to screen for OMPs in water with respect to risk assessment or attenuating them from water sources using an appropriate water treatment process. Within the framework of this study, the present thesis focused on oxidative treatment with ozone and advanced oxidation processes (AOP), developing QSAR models for ozone and AOP. In addition, a decision support system (DSS), based on multi-criteria analysis (MCA), was developed to aid in the choice of an appropriate water treatment process for OMPs removal.

1.3. OXIDATION BY OZONE

The application of ozone in drinking water treatment is widespread throughout the world. The main use of ozone in water treatment are disinfection and oxidation (e.g., taste and odor control, decoloration, destruction of OMPs, etc.) or a combination of both [10]. Similar to other disinfectants for water treatment (e.g. chlorine or chlorine dioxide), ozone is unstable in water and decomposes into hydroxyl radicals ($^{\bullet}$OH) which are the strongest oxidants in water and result in advanced oxidation process (AOP) [11]. Thus, ozonation processes always involve the two species: ozone and $^{\bullet}$OH radicals. However, for different applications of ozone the two species are of differing importance. While disinfection occurs dominantly through ozone, oxidation processes may occur through oxidants, ozone and $^{\bullet}$OH radicals. In conjunction with the benefits of disinfection and oxidation, undesired by-products can be formed from the reaction of ozone and $^{\bullet}$OH radicals with water matrix components. The stability of dissolved ozone is affected by pH, UV, ozone concentration and the concentration of the scavengers. The ozonation reaction is second-order, i.e., first order in ozone and first order in OMP [12]. This reaction is governed by a reaction rate constant (k_{O3}/k_{OH}) with the units, $LMol^{-1}sec^{-1}$.

Ozone reacts with the OMPs either through a ring formation, ozone insertion or electron transfer. In case of AOP, since it involves $^{\bullet}OH$ radicals, it is radical-based, diffusion controlled reaction and proceeds by three steps: initiation, propagation and termination. Thus, the oxidation of an OMP can be predicted by using change in ozone concentration or the second-order rate constants (k_{O3}/k_{OH}) or the percent-removals (%) of the OMPs.

1.4. QSAR – A NOVEL METHOD

As discussed earlier, the analytical methods used to assess elimination of OMPs are not readily available and are time consuming. The OMPs are usually classified based on their origin/source or their ecological/health effects. A more rational way to classify them, however, would be based on the compound physical/chemical properties which can serve as indices for treatment process selection and performance. The Structure Activity Relationship (SAR) which is widely used in medicinal chemistry that correlates the structure of the medicine to the activity intended has found its application in water treatment as well. Such a classification permits a QSAR approach to relationally link compound properties (structure) to treatment process attributes and/or conditions (activity).

QSAR has found its way into the practice of agro chemistry, the pharmaceutical chemistry, toxicology, and other facets of chemistry. Its staying power may be attributed to the strength of its initial hypothesis that activity is a function of structure as described by electronic attributes, hydrophobicity, and steric properties as well as the rapid and extensive development in methodologies and computational techniques that have ensued to delineate and refine the many variables and approaches that it involves [13].

The overall goals of QSAR retain the original essence of the science behind the water-treatment or biological processes and remain focused on the predictive ability of the approach and its receptiveness to mechanistic interpretation [14].

1.5. RESEARCH OBJECTIVES

The objectives of this research thesis are:

- Identify and understand the physical/chemical properties of the OMPs that influence ozonation/oxidation and thereby have a greater understanding of OMPs ozonation/oxidation in water

- Identify and develop the type of QSAR models which help to understand ozonation/advanced oxidation in a more comprehensive and transparent manner

- Identify the best water treatment processes for OMPs removal and the most important socio-economic-technical and sustainability criteria

- Propose hybrids of natural and advanced water treatment processes for OMPs removal for a more sustainable and efficient multi-barrier water treatment approach

1.6. ORGANIZATION OF THESIS

The thesis contains five chapters. The main chapters (*Chapters 2 - 5*) are based on peer-reviewed scientific journal publications and conference presentations.

Chapter 2 deals with QSAR models to predict the percent removal of OMPs by ozonation and AOP for four different river water matrices. Individual models for ozonation and AOP were built for each river water study and finally a global model based on neural networks combining all river waters was developed. The individual models were based on multi-linear regression in order to facilitate better understanding. Compound properties such as energy gap between orbitals and mean oxidation number were studied to understand their influence on ozonation. Internal database was used to develop and validate the QSAR models. Subsequently, validation with an external database was also performed. The models were also defined by a set of boundary conditions such as pH, alkalinity, dissolved organic carbon, etc.

Considering the good modelling results obtained in *Chapter 2, Chapter 3* deals with the development of an extended QSAR model which predicts the ozone and hydroxyl radical rate constants (k_{O3}, k_{OH}). These models were developed for 123 OMPs. The classical statistic techniques used to develop a QSAR model such as: correlation analysis, principal component analysis and multi-linear regression were used. Also, the influence of unsaturation in a compound, ionisation potential and several properties which influence ozonation were studied and the models were validated. Additionally, the classification of OMPs was carried out and a comparison study with the 'Hammett-Taft substituents constants' based QSAR models was carried out.

Chapter 4 describes the DSS based on MCA used to dynamically evaluate the different water treatment processes and several socio-economic-technical and sustainability criteria for OMPs removal. Considering the good results obtained from QSAR models in *Chapter 2* and *Chapter 3* efforts were made to perform MCA studies and create a DSS purely based on computational aspects. Moreover, the chapter focuses on DSS based on pilot scale experimental values. Based on the MCA studies, the natural systems and oxidation performed well for the OMPs removal. It was also observed that OMPs such as carbamazepine, oxybenzone, diazepam and dilantin were better removed by ozonation than natural treatment processes. In addition, among the membrane processes, nanofiltration (NF) can be preferred over reverse-osmosis (RO) for OMPs removal due to similar performance.

Based on the conclusions of the DSS study in *Chapter 4*, wherein different water treatment processes and criteria were considered for OMPs removal, *Chapter 5* presents a new concept in multi-barrier treatment processes with synergies in which two coupled processes can function as a hybrid process. The hybrid processes include a natural treatment process (artificial recharge and recovery) coupled with an advanced process (oxidation, membrane, and adsorption). The study resulted in quadrant plots based on experimental percent removal of OMPs. The hybrids can help to minimize the usage of resources by advanced treatment processes and enhance the usage of the natural treatment processes thereby resulting in a more sustainable approach.

REFERENCES

1. IMS, *IMS Health Market Prognosis* (http://www.imshealth.com/portal/site/ims/). March 2011.

2. Chatzitakis, A., Berberidou, C., Paspaltsis, I., Kyriakou, G., Sklaviadis, T., Poulios, I., *Photocatalytic degradation and drug activity reduction of Chloramphenicol.* Water Research, 2008. **42**(1): p. 386-394.

3. Oaks, J.L., Gilbert, M., Virani, M.Z., Watson, R.T., Meteyer, C.U., Rideout, B.A., Shivaprasad, H., Ahmed, S., Chaudhry, M.J.I., Arshad, M., *Diclofenac residues as the cause of vulture population decline in Pakistan.* Nature, 2004. **427**(6975): p. 630-633.

4. Jobling, S.,Tyler, C.R., *Endocrine disruption in wild freshwater fish.* Pure and Applied Chemistry,, 2003. **75**(11-12): p. 2219-2234.

5. Bergman, Å., Heindel, J.J., Jobling, S., Kidd, K., Zoeller, T.R., *State of the science of endocrine disrupting chemicals 2012: summary for decision-makers.* 2013.

6. Benotti, M.J., Trenholm, R.A., Vanderford, B.J., Holady, J.C., Stanford, B.D., Snyder, S.A., *Pharmaceuticals and Endocrine Disrupting Compounds in U.S. Drinking Water.* Environmental Science and Technology, 2008. **43**(3): p. 597-603.

7. Jones, O.A., Lester, J.N., Voulvoulis, N., *Pharmaceuticals: a threat to drinking water?* Trends in Biotechnology, 2005. **23**(4): p. 163-167.

8. Focazio, M.J., Kolpin, D.W., Barnes, K.K., Furlong, E.T., Meyer, M.T., Zaugg, S.D., Barber, L.B., Thurman, M.E., *A national reconnaissance for pharmaceuticals and other organic wastewater contaminants in the United States — II) Untreated drinking water sources.* Science of The Total Environment, 2008. **402**(2–3): p. 201-216.

9. WHO, *Pharmaceuticals in drinking-water*. 2012. p. 52.

10. Camel, V.,Bermond, A., *The use of ozone and associated oxidation processes in drinking water treatment.* Water Research, 1998. **32**(11): p. 3208-3222.

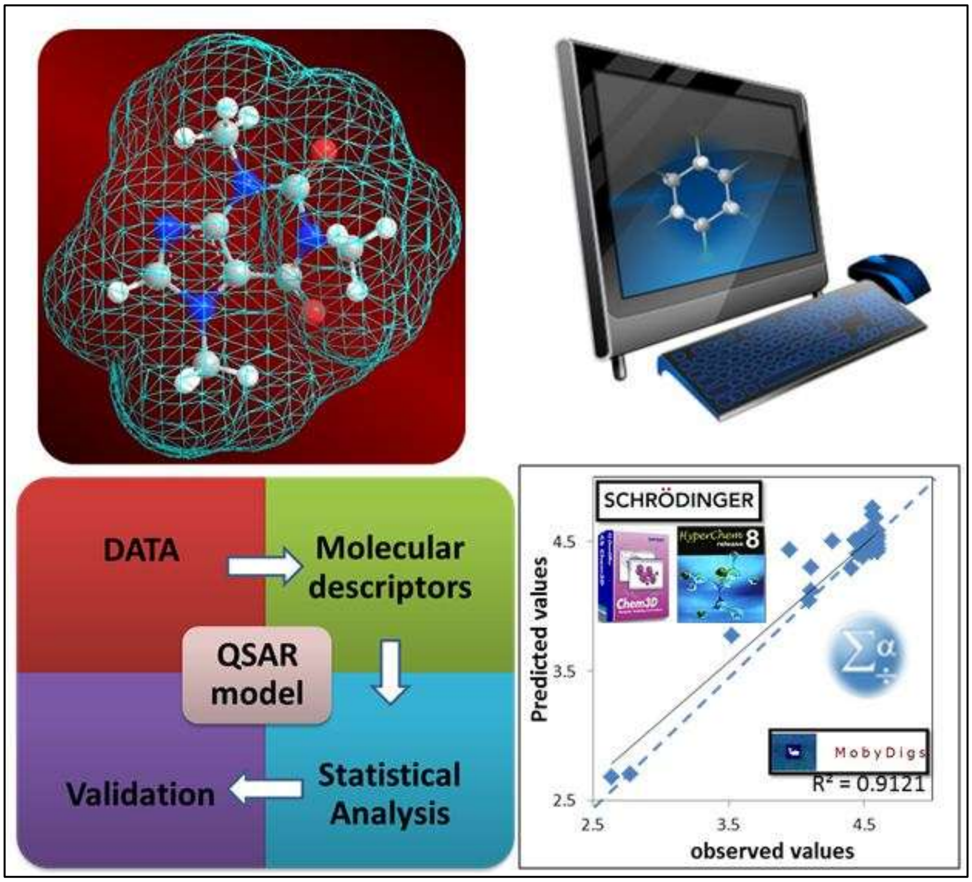
11. Staehelin, J.,Hoigne, J., *Decomposition of ozone in water in the presence of organic solutes acting as promoters and inhibitors of radical chain reactions.* Environmental Science & Technology, 1985. **19**(12): p. 1206-1213.

12. Tomiyasu, H., Fukutomi, H., Gordon, G., *Kinetics and mechanism of ozone decomposition in basic aqueous solution.* Inorganic Chemistry, 1985. **24**(19): p. 2962-2966.

13. Benigni, R., *Quantitative structure-activity relationship (QSAR) models of mutagens and carcinogens*. 2003: CRC Press.

14. Abraham, M.H., Hassanisadi, M., Jalali-Heravi, M., Ghafourian, T., Cain, W.S., Cometto-Muñiz, J.E., *Draize rabbit eye test compatibility with eye irritation thresholds in humans: a quantitative structure-activity relationship analysis.* Toxicological Sciences, 2003. **76**(2): p. 384-391.

2. QSAR MODELS FOR THE REMOVAL OF ORGANIC MICROPOLLUTANTS IN FOUR DIFFERENT RIVER WATER MATRICES



HIGHLIGHTS

- Percent – removal of OMPs by ozonation/advanced oxidation was predicted
- QSAR defining properties: Orbitals energy gap, oxygen/carbon, electron affinity
- Regression based localized QSAR models for individual river water matrices
- Neural network based global QSAR models compiling all river waters data

This chapter is based on the following published article:

Sairam Sudhakaran, James Calvin, Gary Amy. QSAR models for the removal of organic micropollutants in four different river water matrices. *Chemosphere* 87 (2012) 144-150

Sairam Sudhakaran, James Calvin, Gary Amy, QSAR models for ozonation in natural water matrix, European water and wastewater management (EWWM)-2011 (London)

2.1. INTRODUCTION

As discussed in *Chapter 1*, the OMPs are ubiquitous and are prioritized as a list of pollutants in water. There are several candidate water treatment processes (adsorption, membrane separation, river bank filtration) to eliminate the OMPs. Adsorption processes are less efficient with polar compounds. Reverse Osmosis (RO) is a very efficient water treatment process but there are problems of brine solution and expense. . River bank filtration, a sustainable process governed by biodegradation, is less efficient with respect to non-biodegradable compounds (persistent organic micropollutants) such as primidone and atrazine. Ozonation is a good option to remove OMPs since ozonation exhibits selectivity towards certain organic compounds and easily transforms them [1]. Advanced oxidation processes (AOP), exploiting hydroxyl radical ($^{\bullet}$OH) oxidation, are generally considered less selective and hence may oxidize a wider range of compounds.

The analytical methods used to detect them are complicated, expensive, and in certain cases, time-consuming. Predictive models, Quantitative Structure Activity/Property Relationship (QSAR/QSPR) models, are a rapid and cost-effective alternative to experimental evaluation. The number of QSAR articles published in the scientific literature in water-related sciences is constantly increasing. QSAR models are recognized by government regulatory bodies as a method to screen toxic chemicals. Biowin, a software tool that predicts the biodegradability of toxic compounds in water, has its basis in QSAR models [2]. With the increasing growth of reliable software, it has become relatively easy to compute the important properties related to micropollutants. QSAR models are also used to study reaction mechanisms and degradation pathways of micropollutants [3]. QSAR models use relevant molecular physico-chemical properties (molecular descriptors) to predict important treatment responses (e.g., rate constants) which can serve as indices for water treatment process selection and performance assessment. Models have also been developed for membrane separation [4, 5], adsorption [6], and ozonation [7].

Ozone is a strong and selective electrophile. It reacts with double bonds, activated aromatic rings, and specific ring atoms. In general, electron-pumping groups on a benzene ring enhance ozonation and electron-withdrawing groups (nitro, halogens) decrease ozonation. However, apart from the nature of the substituents, other parameters such as steric factors and connectivity between the aromatic rings also need to be considered. Ozone is not very stable in water. It is degraded into $^{\bullet}OH$ radicals which are less selective than ozone, but increase the oxidation efficiency since the reactions are diffusion controlled [1]. Energy of the highest occupied molecular orbital (E_{HOMO}), energy of the lowest unoccupied molecular orbital (E_{LUMO}), and the difference between them, (E_{LUMO} - E_{HOMO}), influence ozonation at the molecular orbital level [7-12]

In this work, QSAR models were developed for the removal of OMPs from water. The molecular descriptors ranged from one-dimensional (atom counts) to three-dimensional (quantum-chemical). The choice of appropriate quantum-chemical descriptors is important because they are relevant to ozonation mechanisms, and can be easily obtained by computation.

2.2. METHODOLOGY

2.2.1. Data Sets

The datasets were taken from bench-scale studies published in a report [13]. The studies were done on Colorado River (CRW), Passaic River (PRW), and Ohio River (ORW) water matrices and synthetic water prepared from Suwannee River (SRW), natural organic matter (NOM) isolate. The model water was prepared by adding Suwannee River Reverse Osmosis (RO) isolated natural organic matter (NOM), purchased from International Humic Substances Society (St. Paul, MN, USA), to de-ionized water with sodium bicarbonate added as a pH buffer. The relevant water quality parameters for the different water matrices and the ozone and peroxide dosages are shown in *Table 2.1*.

Table 2.1: Physical/Chemical conditions for the four river water matrices

Water	Code	pH	DOC (mg/L)	Alkalinity (mg/L as $CaCO_3$)	Dosage (mg/L)		Time (mins)
					O_3	H_2O_2	
Colorado	CRW	8.2	2.5	140	2.5	0.0625	5
Passaic	PRW	6.8	3.4	52	3	0.075	5
Suwannee	SRW	7.5	4.0	400	4*	0.2	5
Ohio	ORW	7.9	3.5	79	3.5	0.0875	5

*8mg/L of O_3 was used for AOP studies

Two studies were performed: one based on an ozone dosage and the other on ozone and hydrogen peroxide dosage, representing an AOP process. The micropollutants were measured at ng/L levels and their percent-removals were determined. The dataset consisting of the percent removals of the OMPs for the different river water matrices are shown in *Table 2.2*. The chemical structures of the OMPs are available in *Appendix A1*.

Table 2.2: Percent removal of OMPs by oxidation from four river water matrices

OMP	CRW		ORW		PRW		SRW	
	O_3	AOP	O_3	AOP	O_3	AOP	O_3	AOP
Acetaminophen	96	96	97	97	96	96	95	95
Androstenedione	99	99	98	98	78	95	98	98
Atrazine	45	52	71	80	17	24	96	96
Benzopyrene	71	71	91	91	67	67	95	91
Caffeine	99	99	99	99	83	87	95	95
Carbamazepine	98	98	98	98	99	99	98	97
DDT	57	61	58	71	35	27	80	85
DEET	76	84	95	96	63	68	99	99
Diazepam	79	86	96	96	66	69	96	99
Diclofenac	96	96	97	97	86	86	93	95
Dilantin	82	89	98	98	72	73	93	95
Erythromycin	96	96	98	98	96	96	no data	no data
Estradiol	99	99	99	99	99	99	97	97
Estriol	99	99	no data	no data	99	99	no data	no data
Estrone	99	99	no data	no data	99	99	no data	no data
Ethinyl estradiol	99	99	99	99	99	99	98	98
Fluorene	94	95	94	94	83	83	no data	no data
Fluoxetine	98	98	92	97	82	81	93	95
Galaxolide	87	89	91	91	76	77	89	89
Gemfibrozil	99	99	98	98	99	99	96	96
Hydrocodone	98	98	no data	no data	99	99	no data	no data
Ibuprofen	82	88	94	96	56	60	97	99
Iopromide	46	60	75	86	52	57	92	92
Lindane	4	14	24	26	1	3	25	30
Meprobamate	50	61	84	85	40	43	97	96
Metolachlor	80	86	92	94	73	82	86	86
Musk ketone	30	34	50	68	20	3	76	75
Naproxen	94	94	95	95	92	92	95	95
Oxybenzone	97	97	97	97	97	97	95	96
Pentoxyfylline	98	98	98	98	86	90	95	93
Progesterone	98	99	97	97	83	96	99	97
Sulfamethoxazole	97	97	95	95	79	79	80	85
TCEP	13	16	5	16	1	1	1	1
Testosterone	99	99	98	98	83	96	99	99
Triclosan	79	82	97	97	98	98	91	96
Trimethoprim	99	99	99	99	99	99	99	96
Range	95	85	94	83	98	98	98	98
Mean	81.44	84.06	86.94	89.21	73.69	75.5	87.83	88.35

2.2.2. Molecular Descriptors Computation

The descriptors chosen for this study ranged from constitutional to quantum-chemical. Some of the constitutional descriptors are the number of halogens; hydrogen, nitrogen and oxygen normalized to carbon; mean oxidation number (MON); and double bond equivalence (DBE). The geometric descriptors related to surface area include solvent accessible surface area (SASA), pi-surface area (PISA), polar surface area (PSA), ionization potential (IP), and electron-affinity (EA), and were computed using Qikprop, version 3.2, Schrodinger software. The quantum-chemical descriptor related to molecular orbitals, the difference in energy between lowest unoccupied and highest occupied molecular orbital (E_{LUMO} - E_{HOMO}), was computed using ChemBio3D Ultra 11.0. More than 30 descriptors which influence ozonation were computed for the model(s) development (*Table 2.3*). The molecular structures were geometrically optimized by the semi-empirical method MMFF94 (Merck Molecular Force Field) [14] in ChemBio3D Ultra 11.0. The optimizations were necessary to compute the quantum chemical and geometric descriptors. The molecular structures of the compounds were obtained using SMILES (simplified molecular input line entry specification) notation and the 2-D structures were converted to 3-D using ChemBio3D Ultra 11.0. The possible QSAR model-defining molecular descriptors along with their corresponding values are listed in *Table 2.4.*

2.2.3. QSAR model and validation

SPSS version 17.0 was used to build the QSAR models. In some models, the percent-removal (% rmvl) of organic micropollutants was transformed to natural logarithm (lnrmvl) in order to handle extreme values and symmetrize the responses [6, 9, 15]

Table 2.3: List of molecular descriptors involved in the QSAR building process

Descriptor/Property	**Description**
Quantum	
E_{HOMO}, E_{LUMO}	energy of the highest occupied and lowest unoccupied molecular orbitals in eV
E_{LUMO} - E_{HOMO}	energy difference between the lowest unoccupied and highest occupied molecular orbital in eV
Topological	
MW	molecular weight in grams
dipole	dipole moment (D, debye)
SASA	solvent accessible surface area using a probe with 1.4 Angstroms radius (A^2)
FOSA,FISA, PISA,WPSA	hydrophobic, hydrophilic, Pi, weakly polar components of SASA in square angstroms (A^2)
volume	total solvent-accessible volume in cubic angstroms computed using a probe with 1.4 Angstroms radius
donorHB, accptHB	number of hydrogen bonds donated or accepted by the solute to/from water
polrz	polarizability in cubic angstroms
logPo/w	octanol/water partition co-efficient
logS	water solubility in moles per cubic decimeters
IP, EA	ionisation potential and electron-affinity in eV
Atom counts	
#rotor	# of non-trivial, non-hindered rotatable bonds
#rtvFG	# of reactive functional groups(amides, hetero atoms, azo, diazo, azide)
#ringatoms,#in34, #in56	# of ring atoms, number of atoms in 3- or 4-; 5-or 6- membered rings
#noncon	# of ring atoms unable to form conjugated aromatic systems (sp^3 carbon)
#nonHatm	# of heavy atoms (non-hydrogen atoms)
#X	# of halogens
#C=C	# of carbon-carbon double bonds
H/C, N/C, O/C	hydrogen, nitrogen, oxygen to carbon ratios
#Arom	# of aromatic rings
MON	mean oxidation number
DBE	double bond equivalence

Table 2.4: OMPs along with the QSAR – defining molecular descriptor values

molecule	E_L-E_H	EA	#ringatoms	MON	#X	#in56	O/C
Acetaminophene	8.533	-0.22	6	-0.25	0	6	0.25
Androstenedione	8.6	0.035	10	-1.15	0	10	0.11
Atrazine	9.00	-0.218	6	0.25	1	6	0.00
Benzopyrene	8.50	1.289	20	-0.60	0	20	0.00
Caffeine	8.41	0.443	9	0.75	0	9	0.25
Carbamazepine	8.1	0.772	15	-0.26	0	12	0.07
DDT	9.1	0.525	12	-0.28	5	12	0.00
DEET	8.5	0	6	-1.00	0	6	0.08
Diazepam	8.421	0.859	17	-0.25	1	12	0.06
Diclofenac	8.381	0.355	12	-0.14	2	12	0.14
Dilantin	8.5	0.329	17	-0.13	0	17	0.13
Erythromycin	8.746	-0.301	26	-1.02	0	12	0.35
Estradiol	8.5	-0.497	10	-1.11	0	10	0.11
Estriol	8.5	-0.5	10	-1.00	0	10	0.17
Estrone	8.5	-0.418	6	-1.00	0	6	0.11
Ethinyl estradiol	8.5	-0.504	10	-1.00	0	10	0.10
Fluorene	8.508	0.299	13	-0.77	0	13	0.00
Fluoxetine	8.075	0.048	12	-0.58	3	12	0.06
Galaxolide	9	-0.495	13	-1.33	0	13	0.06
Gemfibrozil	8.5	-0.38	6	-1.06	0	6	0.20
Hydrocodone	8.951	0.494	16	-0.66	0	16	0.17
Ibuprofen	9	-0.36	6	-1.07	0	6	0.15
Iopromide	8.606	1.759	6	0.23	3	6	0.44
Lindane	11	0.02	6	0.00	6	6	0.00
Meprobamate	9	-0.864	0	-0.44	0	0	0.44
Metolachlor	9	-0.237	19	-0.93	1	10	0.13
Musk ketone	9.272	1.659	6	-0.14	0	6	0.36
Naproxen	8.15	0.54	10	-0.57	0	10	0.21
Oxybenzone	8.711	0.67	12	-0.42	0	12	0.21
Pentoxyfylline	8.391	0.321	9	0.00	0	9	0.23
Progesterone	8.402	0.049	17	-1.23	0	17	0.10
Sulfamethoxazole	8.626	0.592	11	-0.20	0	11	0.30
TCEP	11	0.474	0	-1.33	3	0	0.67
Testosterone	8.893	0.118	17	-1.26	0	17	0.11
Triclosan	8.669	0.621	12	0.00	3	12	0.17
Trimethoprim	8.261	0.127	12	0.00	0	12	0.21

Model selection techniques using multiple linear regression (MLR) and artificial neural networks (ANN) [4, 14, 16] were used to develop the model using the molecular descriptors to predict percent removal, or lnrmvl. The analysis of variance (ANOVA) method was used to test the model parameters. The ANOVA method is used to study the influence of the independent variables on the dependent variables. The result of the model building process is an explicit function relating the set of predictors to the dependent variable. The model selection process involved iterative model fitting using MLR with intermediate model analysis to assess the quality of the intermediate model and to determine the next step in the iterative process. At each stage, the fit of the model, the change in the fit of the model produced by a reduction in the number of predictors, the model r^2, and the significance of the individual variables were used to direct the process. A p-value of 0.05 was used to determine the significance of potential predictors in the model building process. Variables with high p-values were sequentially removed in order to develop a more parsimonious model. This process was performed until further reduction in the model significantly reduced the predictive value of the equation. r^2, the proportion of variability explained by the model, was also computed. r^2 values range from 0 to1 and a high r^2 indicates that the model explains a large proportion of the variability in the response.

A predictive model based on artificial neural networks (ANN) was also developed compiling data from all the river water matrices. A feed forward architecture, wherein the connections in the network flow forward from the input layer to the output layer without any feedback loops was employed, and multi-layer perceptron (MLP) network procedure was used [17]. The model defining co-variants/descriptors were the same as used in AOP and ozone QSAR models. The co-variants were rescaled using the standardized method where, the value is subtracted by the mean and divided by the standard deviation.

Validation of the QSAR models is performed to check the robustness and predictive power of the models. The validation of the final QSAR models was performed using leave-one-out (LOO) cross-validation and bootstrap methods using MobyDigs, version 11.0 (Talete, Milan). LOO internal cross-validation is well-suited for a small dataset [18]. With LOO, one point of the dataset is removed and a QSAR model is built with the remaining data and the value of the omitted point is predicted. This process is carried out for every data point, yielding Q^2_{LOO}:

$$Q^2_{LOO} = 1 - \frac{PRESS}{TSS};$$

TSS: the total corrected sum of squares, PRESS: predicted error sum of squares, sum of squared differences between the observed and estimated response from the model built without that observation

Q^2_{LOO} increases only when the added descriptors influence the QSAR equation. A Q^2_{LOO} value greater than 0.5 is generally regarded as good [19]. Additionally, an external validation was performed for the CRW ozone data since external data with similar experimental conditions, were available only for direct ozonation [7]. The data for external validation were obtained from a literature study [7].

In case of the bootstrap method, theoretical samples from the same empirical distribution are created using the information from the original sample. The validation is performed by randomly generating training sets with sample repetitions and then evaluating the predicted responses of the samples not included in the training set. The model is calculated with the training set and the responses are predicted with the evaluation set. All the squared differences between the true response and the predicted response of the objects of the evaluation set are gathered in PRESS (predictive error sum of squares). This procedure of building training sets and evaluation sets is repeated 5000 times and the PRESS are summed up and the average predictive power is calculated (Q^2_{BOOT}) [20].

2.3. RESULTS AND DISCUSSION

2.3.1. QSAR models: CRW, ORW, PRW and SRW

The AOP – QSAR and O_3 – QSAR models for CRW, ORW, PRW and SRW along with the associated statistical indices such as r^2 and q^2 are shown in *Table 2.5*. The QSAR models for both AOP and O_3 had a high goodness of fit, $r^2 > 0.8$, and good internal validation, (Q^2_{LOO} and $Q^2_{BOOT} > 0.5$) which indicates that the models are very robust and predictive. An external validation was performed for the CRW - O_3 data. The external dataset consisted of 22 compounds as shown in *Table 2.6*. The external validation resulted in a goodness of fit, $r^2 = 0.77$. Graphs of observed vs predicted values for the QSAR models and external validation are shown in *Figure 2.1*.

The molecular descriptors which influenced the QSAR models were: $E_{LUMO} – E_{HOMO}$, EA, #ring atoms, MON, #X, #in56, O/C.

The frontier orbital energies are important in radical reactions such as ozonation. The energy of the highest occupied molecular orbital (E_{HOMO}) indicates the susceptibility of a molecule to be attacked by an electrophile such as ozone [21]. Energy of the lowest unoccupied molecular orbital (E_{LUMO}) indicates the susceptibility of a molecule to be attacked by a nucleophile [21]. The energy gap, E_{LUMO} - E_{HOMO}, is the energy difference between the lowest unoccupied and highest occupied molecular orbitals of the micropollutant (*Figure 2.2*). E_{LUMO} - E_{HOMO} is a good index for radical reaction potential in a hydroxyl radical system. Compounds with large gap (E_{LUMO} - E_{HOMO}), have low chemical reactivity towards ozone [8, 12, 22, 23]

Electron-Affinity (EA) is a measure of the affinity of the molecule towards electrons. In the case of ozonation, ozone is an electrophile and has a high electron-affinity. OMPs with low electron-affinity enhance ozonation since there is no competition between ozone and the OMPs for the electrons.

Number of ring atoms (#ring atoms) and number of ring atoms in 5 or 6 membered ring (#in56) indicates the presence of carbon based ring systems. The presence of more ring structures (aromatic, double-bond) indicates high electron-density and hence can easily undergo ozonation [1].

The mean oxidation number (MON) indicates the oxidation state of the OMPs. The oxidation states for carbon range from -4 to 4. Compounds with low MON can easily undergo ozonation since the carbon is in a more reduced state [24].

The number of halogens (#X) in a molecule is influential towards ozonation process, since halogens are electrophiles. Halogens such as chlorine and fluorine when attached to the aromatic/aliphatic system withdraw the electrons towards them and make the OMPs electron-deficient for ozone, an electrophile. Overall, #X decreases ozonation capacity.

The oxygen to carbon ratio (O/C) indicates oxidizability of the compound. Compounds with high O/C values have low reactivity with ozone as compared to compounds with low O/C.

Table 2.5: QSAR models for the four – different rivers: CRW, ORW, PRW, SRW

CRW	AOP	lnrml (AOP) = 9.77 - 0.63 (E_{LUMO}-E_{HOMO}) - 0.194 (EA) +0.02 (#ringatoms) r^2 = 0.902, Q^2_{LOO} = 0.868, Q^2_{BOOT} = 0.819
	O_3	lnrml (O_3) = 12.45 - 0.95 (E_{LUMO}-E_{HOMO}) – 0.32 (MON) r^2 = 0.866, Q^2_{LOO} =0.792, Q^2_{BOOT} = 0.736
ORW	AOP	%rml (AOP) = 318.23 – 26.52 (E_{LUMO} - E_{HOMO}) – 6.11 (EA) + 0.41 (#ringatoms) r^2 = 0.922, Q^2_{LOO} = 0.897, Q^2_{BOOT} = 0.838
	O_3	%rml (O_3) = 310.31 – 25.91 (E_{LUMO} - E_{HOMO}) – 8.64 (EA) – 2.11 (#X) +0.66 (#ringatoms) r^2 = 0.915, Q^2_{LOO} = 0.853, Q^2_{BOOT} = 0.801
PRW	AOP	lnrml (AOP) = 15.17 – 1.33 (E_{LUMO} - E_{HOMO}) - 0.56 (EA) +0.06 (#in56) r^2 = 0.862, Q^2_{LOO} = 0.790, Q^2_{BOOT} = 0.740
	O_3	lnrml (O_3) = 18.15 - 1.63 (E_{LUMO} - E_{HOMO}) – 0.33 (MON) r^2 = 0.887, Q^2_{LOO} = 0.851, Q^2_{BOOT} = 0.772
SRW	AOP	%rml (AOP) = 286.67 – 22.41 (E_{LUMO} - E_{HOMO}) – 52.68 (O/C) r^2 = 0.854, Q^2_{LOO} = 0.702, Q^2_{BOOT} = 0.603
	O_3	%rml (O_3) = 345.54 – 29.18 (E_{LUMO} - E_{HOMO}) – 5.59 (EA) r^2 = 0.862, Q^2_{LOO} = 0.78, Q^2_{BOOT} = 0.65

2.3.2. Statistical significance of the QSAR-descriptors

The F-statistic and the p-values/significance levels are used to test the null - hypothesis. The null- hypothesis proposes that no statistical significance exists in a set of given observations, therefore, the null hypothesis needs to be rejected for a linear relationship to exist. The F-statistic is the ratio of the two mean squares. When the F-value is large and the p-value/significance is low, the null-hypothesis can be rejected. A small p-value/significance level indicates that the results are not due to random chance. The p-values/significance for the QSAR descriptors were less than 0.05, hence rejecting

Table 2.6: External validation dataset

Compound	**%rml**	**lnrml**	**predicted**	**E_L-E_H**	**MON**
Acenaphthene	90	4.50	4.86	8.24	-0.83
Acenaphthylene	91	4.51	5.04	8.00	-0.66
Aldrin	65	4.17	3.95	8.96	-0.16
Anthracene	92	4.52	5.18	7.86	-0.71
Benzo(a)anthracene	86	4.45	4.81	8.23	-0.66
Benzo(a)pyrene	75	4.32	5.05	7.96	-0.60
Benzo(b)fluoranthene	84	4.43	4.84	8.18	-0.60
Benzo(k)fluoranthene	80	4.38	5.20	7.81	-0.60
Chrysene	86	4.45	4.77	8.28	-0.66
DDD	73	4.29	3.96	9.05	-0.42
DDE	60	4.09	4.87	8.04	-0.28
Dieldrin	4	1.39	2.91	10.01	0.00
Endrin	29	3.37	3.85	9.02	0.00
Heptachlor	53	3.97	3.84	8.96	0.20
Methoxychlor	89	4.49	4.04	8.99	-0.50
Naphthalene	82	4.41	4.66	8.43	-0.80
Octylphenol-4t	93	4.53	4.07	9.26	-1.42
Phenanthrene	95	4.55	4.85	8.21	-0.71
Pyrene	94	4.54	4.92	8.10	-0.62
α-BHC	12	2.48	2.58	10.34	0.00
β-BHC	11	2.40	2.10	10.85	0.00
δ-BHC	7	1.95	2.44	10.49	0.00

the null hypothesis and indicating a 95% confidence interval in the linear relationship and the F-statistic associated with the different models ranged from 80 to 130 (*Table 2.7*). In order to avoid any redundancy while choosing the QSAR-equation descriptors, the pairwise correlations [14] among the descriptors were evaluated. E_{LUMO} - E_{HOMO} and EA appear in all the equations but there are no intercorrelations (-0.008) between them. #ring atoms and #in56 describe a similar property, i.e., the ring-atoms present in a compound; hence an intercorrelation (0.855) exists between them. However, both #ring atoms and #in56 are not present in the same QSAR-equation (*Table 2.8*). The correlation analyses of the molecular descriptors with the %removals are available in *Appendix B1*

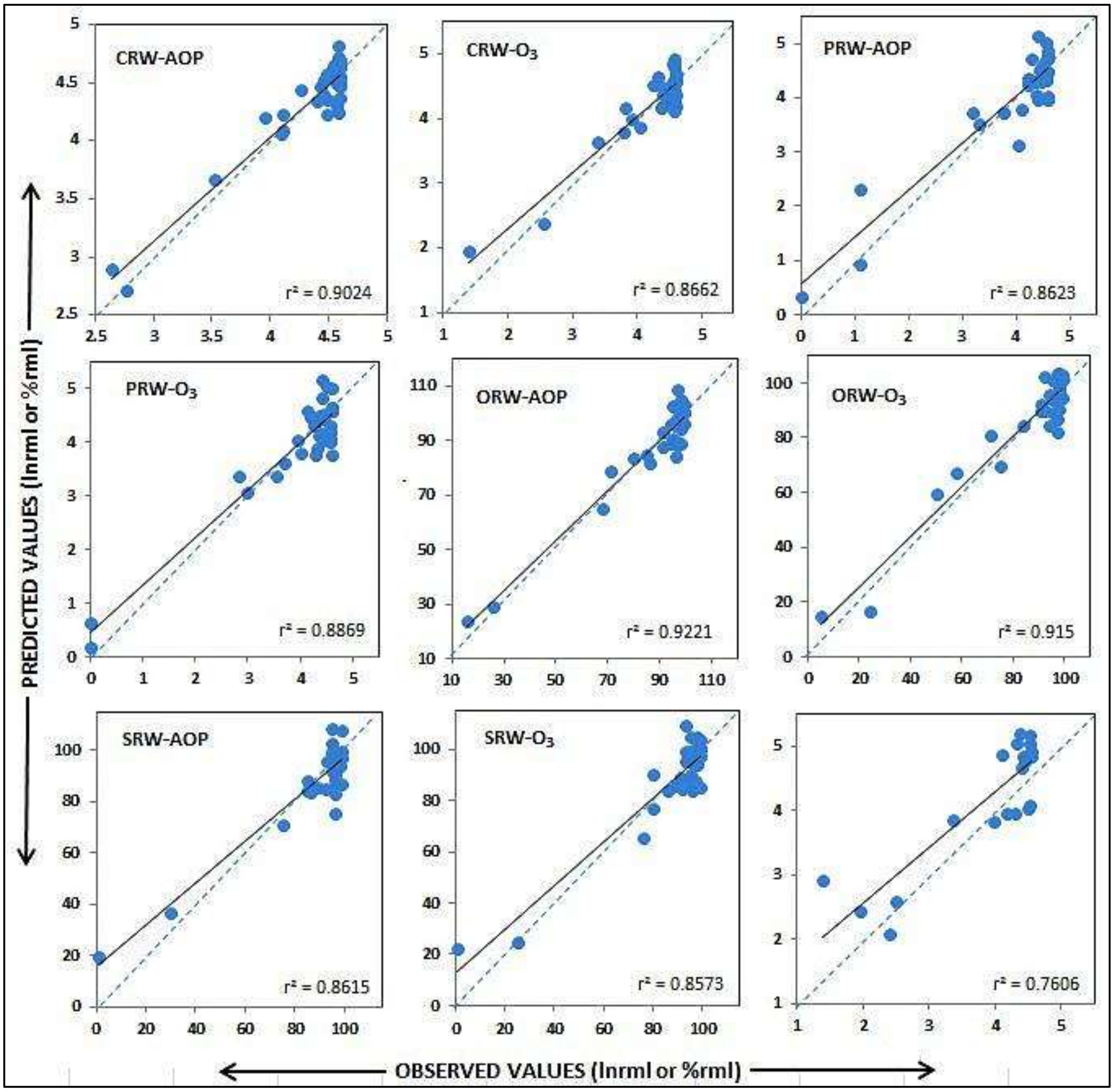


Figure 2.1: QSAR prediction plots for the 4 river water matrices
Graphs of predicted vs observed OMP removals for CRW, PRW, ORW, SRW and external validation for CRW-O3 data (on right). The graphs for ORW-O3, SRW (AOP and O3) are in percent-removals and the others are in lnrml. The bold line is the regression line and the dotted line is the ideal reference line

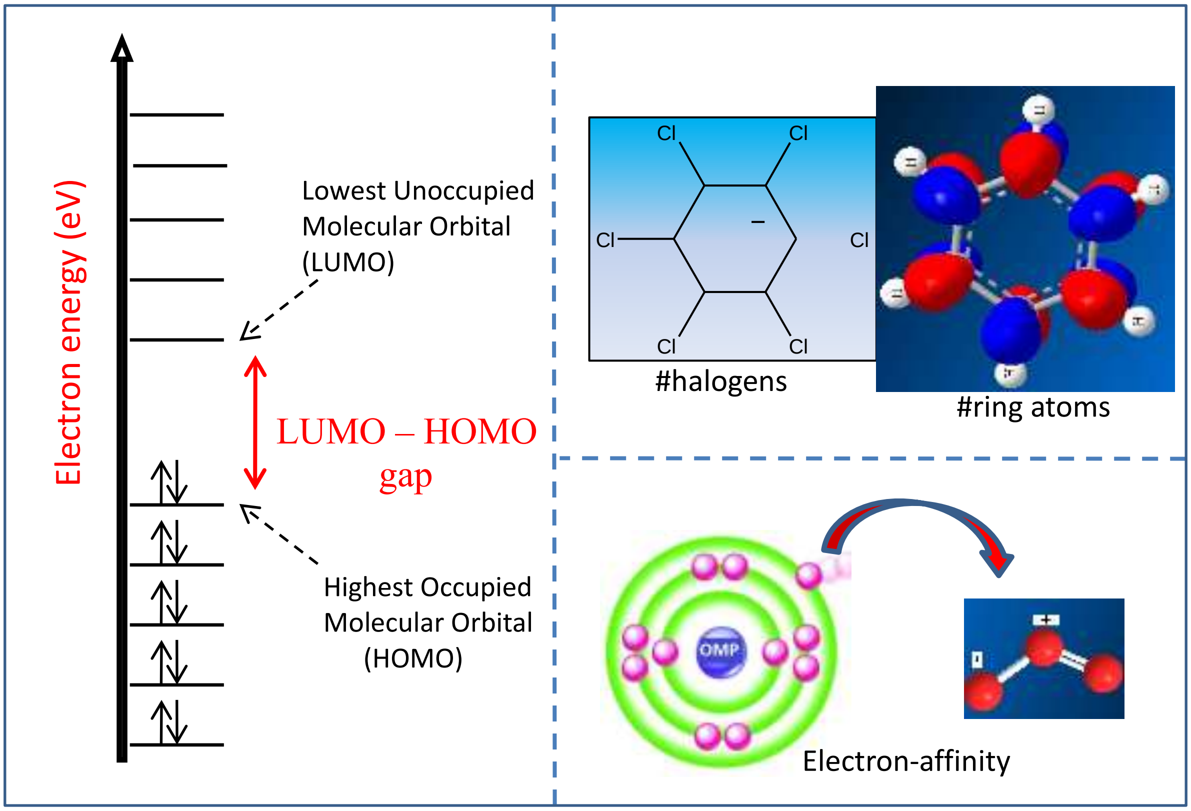


Figure 2.2: Schematic depiction of some QSAR defining molecular descriptors

Table 2.7: p-values/significance levels of the descriptors

Water	**CRW**		**ORW**		**PRW**		**SRW**	
	AOP	**O_3**	**AOP**	**O_3**	**AOP**	**O_3**	**AOP**	**O_3**
E_{LUMO}-E_{HOMO}	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EA	0.000	n.a	0.001	0.000	0.000	n.a	n.a	0.028
#ring atoms	0.003	na	0.043	0.010	n.a	n.a	n.a	n.a
MON	n.a	0.000	n.a	n.a	n.a	0.010	n.a	n.a
#X	n.a	0.001	n.a	0.038	n.a	n.a	n.a	n.a
#in56	n.a	n.a	n.a	n.a	0.003	n.a	n.a	n.a
O/C	n.a	n.a	n.a	n.a	n.a	n.a	0.042	n.a
F-statistic	98.66	67.9	180.3	53.79	64.4	91.02	102.02	120.34

n.a: not applicable

Table 2.8: Pairwise intercorrelations among the descriptors

	E_L-E_H	**EA**	**#ringatoms**	**#in56**	**#X**	**O/C**	**MON**
E_L-E_H	1	-0.008	-0.362	-0.436	0.564	0.301	-0.095
EA		1	0.135	0.233	0.211	0.168	0.479
#ringatoms			1	**0.855**	-0.188	-0.396	-0.128
#in56				1	-0.19	-0.550	-0.066
#X					1	-0.038	0.277
O/C						1	0.048
MON							1

2.3.3. Influence of water quality conditions on OMP removal

Water quality conditions, such as DOC (dissolved organic carbon), alkalinity and pH, influenced the OMPs removal. With an increase in pH and alkalinity, more O_3 is converted into $^{\bullet}OH$ radicals, resulting in AOP and enhanced OMPs removal. In the case of high DOC, $^{\bullet}OH$ radicals are consumed by DOC and ozonation efficiency is decreased. CRW and SRW showed good removals for most of the OMPs due to high pH, alkalinity and low DOC as compared to PRW (low pH, alkalinity and high DOC), ORW was intermediate in response as shown in *Table 2.1*.

2.3.4. ANN-based QSAR models: Compiled data from the four rivers

The data from all four rivers were compiled and QSAR-models for AOP and O_3 were developed by artificial neural networks (ANN). ANN was used as an alternate modeling approach. The dataset was split into three segments: training, test and hold-out sets. The test and the hold-out sets are internal and external validation sets. AOP and O_3 had 73% and 68% training set, 19% test-set, and 8% and 13.2% hold-out set, respectively. Descriptors which had high correlation with % rml were chosen for the ANN analysis. ANN has also been used in QSAR modeling to enhance the proportion of variability [5]. *Figure 2.3* shows the predicted vs observed plots for the ANN-based studies for AOP and

O_3 data. In the AOP-dataset, E_{LUMO} - E_{HOMO} had the highest normalized importance (~100%), followed by #in56 (~39%), DBE (32%) and WPSA (~20%).

In the O_3-dataset, E_{LUMO}- E_{HOMO} had the highest normalized importance (~100%) followed by #ringatoms (~38%), EA (~31%) and #X (~30%).

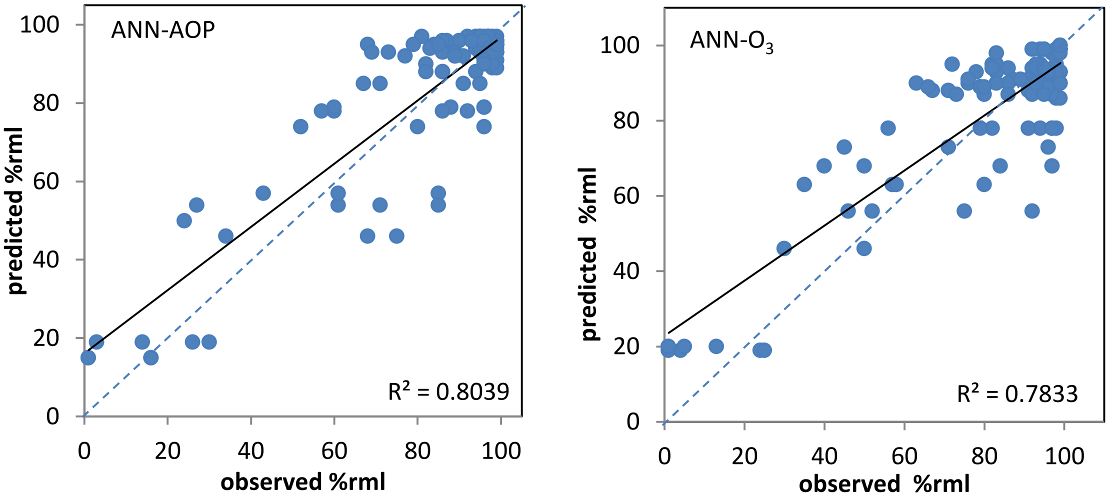


Figure 2.3: ANN based prediction plots for the 4 river water matrices

2.4. CONCLUSIONS

QSAR models focusing on AOP and O_3 for four different river water matrices: CRW, ORW, PRW and SRW; were developed. % rmvl or lnrmvl of OMPs by oxidation were predicted. The energy gap, E_{LUMO} - E_{HOMO}, EA, #X, #in56, O/C, MON were important descriptors in the QSAR equations. E_{LUMO} - E_{HOMO} appeared in all the QSAR models which indicated that it was an important property in understanding ozonation mechanism. Finally, ANN-based models for AOP and O_3 were built compiling data from all the four river water matrices. The proportion of variability, r^2, for all the models were greater than 0.8. The internal and external validations for the models resulted in good predictive powers, Q^2's greater than 0.5. The MLR-based QSAR models are more localized to the individual river waters whereas the ANN-based QSAR models are more global which include data from all the four rivers.

The QSAR models are boundary conditioned to the associated water quality conditions such as pH, DOC, and alkalinity levels. OMPs under similar experimental conditions can be predicted using the QSAR models. Although, the datasets had a large number of percent removals at the higher end (>85%), and very few at the lower end (<50%), good internal and external validation values (>0.5) proved that the models were robust and predictive.

REFERENCES

1. von Gunten, U., *Ozonation of drinking water: Part I. Oxidation kinetics and product formation.* Water Research, 2003. **37**(7): p. 1443-1467.

2. Cronin, M.T., Walker, J.D., Jaworska, J.S., Comber, M.H., Watts, C.D., Worth, A.P., *Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances.* Environmental Health Perspectives, 2003. **111**(10): p. 1376.

3. Sabljic, A., *QSAR models for estimating properties of persistent organic pollutants required in evaluation of their environmental fate and risk.* Chemosphere, 2001. **43**(3): p. 363-375.

4. Yangali-Quintanilla, V., Sadmani, A., McConville, M., Kennedy, M., Amy, G., *A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes.* Water Research, 2010. **44**(2): p. 373-384.

5. Yangali-Quintanilla, V., Verliefde, A., Kim, T.U., Sadmani, A., Kennedy, M., Amy, G., *Artificial neural network models based on QSAR for predicting rejection of neutral organic compounds by polyamide nanofiltration and reverse osmosis membranes.* Journal of Membrane Science, 2009. **342**(1-2): p. 251-262.

6. de Ridder, D.J., Villacorte, L., Verliefde, A.R.D., Verberk, J.Q.J.C., Heijman, S.G.J., Amy, G.L., van Dijk, J.C., *Modeling equilibrium adsorption of organic micropollutants onto activated carbon.* Water Research, 2010. **44**(10): p. 3077-3086.

7. Lei, H.,Snyder, S.A., *3D QSPR models for the removal of trace organic contaminants by ozone and free chlorine.* Water Research, 2007. **41**(18): p. 4051-4060.

8. Karelson, M., Lobanov, V., Katritzky, A., *Quantum-chemical descriptors in QSAR/QSPR studies.* Chemical Reviews, 1996. **96**(3): p. 1027-1044.

9. Kusic, H., Rasulev, B., Leszczynska, D., Leszczynski, J., Koprivanac, N., *Prediction of rate constants for radical degradation of aromatic pollutants in water matrix: A QSAR study.* Chemosphere, 2009. **75**(8): p. 1128-1134.

10. Wang, Y., Chen, J., Li, X., Wang, B., Cai, X., Huang, L., *Predicting rate constants of hydroxyl radical reactions with organic pollutants: Algorithm, validation, applicability domain, and mechanistic interpretation.* Atmospheric Environment, 2009. **43**(5): p. 1131-1135.

11. Pompe, M.,Veber, M., *Prediction of rate constants for the reaction of O3 with different organic compounds.* Atmospheric Environment, 2001. **35**(22): p. 3781-3788.

12. Shiyun, Z., Xuesong, Z., Daotang, L., Weimin, C., *Ozonation of naphthalene sulfonic acids in aqueous solutions: Part II--Relationships of their COD, TOC removal and the frontier orbital energies.* Water Research, 2003. **37**(5): p. 1185-1191.

13. Snyder, S.A., Wert, E.C., Lei, H., Westerhoff, P., Yoon, Y., *Removal of EDCs and Pharmaceuticals in Drinking and Reuse Treatment Processes. Project # [2758]*. 2007, Awwa Research Foundation Denver, Colorado. p. 331.

14. Ott, L.,Longnecker, M., *An Introduction to Statistical Methods and Data Analysis*. 2008: Duxbury Press.

15. Oberg, T., *A QSAR for the hydroxyl radical reaction rate constant: validation, domain of application, and prediction.* Atmospheric Environment, 2005. **39**(12): p. 2189-2200.

16. Gramatica, P., *Principles of QSAR models validation: internal and external.* QSAR & Combinatorial Science, 2007. **26**(5): p. 694-701.

17. Peterson, C.,Soderberg, B., *Artificial neural networks*. 1993: Blackwell Scientific Publications, Oxford.

18. Redding, A.M., Cannon, F.S., Snyder, S.A., Vanderford, B.J., *A QSAR-like analysis of the adsorption of endocrine disrupting compounds, pharmaceuticals, and personal care products on modified activated carbons.* Water Research, 2009. **43**(15): p. 3849-3861.

19. Eriksson, L., Jaworska, J., Worth, A.P., Cronin, M.T.D., McDowell, R.M., Gramatica, P., *Methods for reliability and uncertainty assessment and for applicability evaluations of classification-and regression-based QSARs.* Environmental Health Perspectives, 2003. **111**(10): p. 1361.

20. Efron, B.,Tibshirani, R., *An introduction to the bootstrap*. Vol. 57. 1993: Chapman & Hall/CRC.

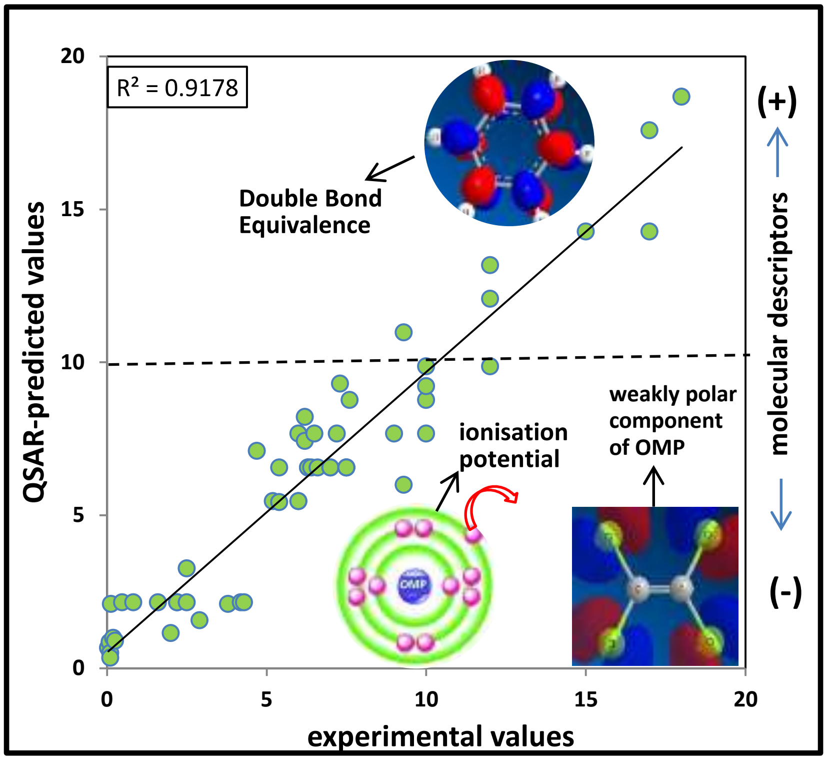
21. Karelson, M., Lobanov, V.S., Katritzky, A.R., *Quantum-chemical descriptors in QSAR/QSPR studies.* Chemical reviews, 1996. **96**(3): p. 1027-1044.

22. Rauk, A., *Orbital interaction theory of organic chemistry*. 2001: John Wiley and Sons.

23. Yan, C., Chen, J., Huang, L., Ding, G., Huang, X., *Linear free energy relationships on rate constants for the gas-phase reactions of hydroxyl radicals with PAHs and PCDD/Fs.* Chemosphere, 2005. **61**(10): p. 1523-1528.

24. Vogel, F., Harf, J., Hug, A., Von Rohr, P.R., *The mean oxidation number of carbon (MOC)--a useful concept for describing oxidation processes.* Water Research, 2000. **34**(10): p. 2689-2702.

3. QSAR MODELS FOR OXIDATION OF ORGANIC MICROPOLLUTANTS IN WATER BASED ON OZONE AND HYDROXYL RATE CONSTANTS AND THEIR CHEMICAL CLASSIFICATION



HIGHLIGHTS

- Ozone and hydroxyl radical rate constants were predicted
- QSAR defining descriptors: Pi-bonds, halogen surface area, ionisation potential
- OMPs structural classification in prediction plots
- Regression based models, validated and mechanistically interpreted

This chapter is based on the published article and conference proceedings:

Sairam Sudhakaran, Gary L. Amy. QSAR models for oxidation of organic micropollutants in water based on ozone and hydroxyl radical rate constants and their chemical classification. *Water Research* 47 (2013) 1111 – 1122

Sairam Sudhakaran, Gary L. Amy., July 2013 (AEESP): QSAR models for oxidation of organic micropollutants in water based on ozone and hydroxyl radical rate constants and their chemical classification. Colorado, USA

3.1. INTRODUCTION

In the last chapter, QSAR models for ozonation and AOP were developed. These models predicted the percent removal of OMPs.

Apart from percent-removal which was used to predict oxidation capacity in the previous chapter, second order rate constants (k_{OH} and k_{O3}) are good indices to estimate the oxidation efficiency, where higher rate constants indicate more rapid oxidation. During the ozonation process, the rate of the reaction is indicated by a second order rate-constant (direct ozone: k_{O3}, AOP: k_{OH}), which is a constant for a given reaction at a particular temperature. The k_{OH} values ($10^9 M^{-1}s^{-1}$) are much higher than the corresponding k_{O3} values ($\sim 10 M^{-1}s^{-1}$) which indicate that $^{\bullet}OH$ radical mediated reactions are faster reactions since they are radical-based and diffusion controlled. The ozonation of OMPs is a second-order type of reaction, i.e., first order with respect to ozone and first order with respect to an OMP. There is plenty of scientific literature focusing on the ozone/hydroxyl radical rate constants which are also compiled in the kinetic data base Radiation Chemistry Data Center of the Notre Dame Radiation Laboratory, available at http://kinetics.nist.gov/solution/. The decomposition of dissolved ozone is highly affected by pH, ozone concentration, and the concentration of various scavengers [1].

As discussed in the previous chapter, the QSAR models are based on the concept that the structure of a molecule influences its properties and are an interdiscipline between chemistry/biology and statistics. There are different statistical approaches in building QSAR models. The most frequently used methods are multiple linear regression (QSAR models developed in the *Chapter 2*), principal component analysis (PCA) and factor analysis, principal regression analysis, partial least squares, discriminant analysis, and neural networks.

There are also 3-D QSAR methods wherein the three-dimensional structure of the molecule and its interaction with its surroundings (solvent, other molecules, etc.) is studied, and comparative molecular field analysis (CoMFA) is the most commonly used approach [2]. While developing a QSAR model there are three components to focus on: dataset, molecular descriptors and statistical technique (*Figure 3.1*). The dataset for the QSAR should be reliable and measured in a consistent manner. The molecular descriptors should be mechanistically related to the predicted property/activity. In the case of descriptors governed by structural conformation such as molecular orbital energies, ionisation potential, etc., proper energy optimization methods must be used. In case of statistical analyses, methods that are simple, transparent and easily interpretable should be the first priority. Transparent models can be easily understood and updated. The transparency decreases with progressing from regression, to partial least squares and finally neural networks, however, the type of dataset also plays a role in choice of the statistical technique. The QSAR model should not be used to make predictions that extrapolate beyond the conditions associated with the model. Finally, validation ensures that the QSAR model can be used for prediction. For this purpose, a certain proportion of the training dataset called the test set (usually up to 50%) can be chosen. The test set should be representative of the complete dataset [3].

QSAR models are also developed based on correlations between rate constants and substituent descriptor constants such as Hammett/Taft constants. In this type of study, emphasis is on compounds with a common parent structure (phenol, amine etc.) and other parts of the molecules are considered as a substituent, and their corresponding Hammett/Taft (σ/σ^*) constants are used and their correlations with the rate constants are studied. These constants indicate the electron-donating and withdrawing properties of the substituents. In the case of complex structures, such as pharmaceuticals, structural approximations are used to compute the Hammett/Taft constants. The structural approximation is based on the premise that inductive/resonance effects of substituent atoms are attenuated with increasing distance from the reaction center. Finally, a linear correlation equation between the rate constants and the substituent descriptors is the QSAR model [4].

In *chapter 2*, the QSAR models based on percent removals were defined by a set of boundary conditions such as pH, DOC and alkalinity. The dataset used for the model was slightly skewed with more OMPs on the lower and upper range (>90% and <40%). Alternatively, the rate constants (k_{OH}, k_{O3}) are not influenced by dissolved organic matter present in water. The apparent rate constants are only affected by the speciation (positive/negative) of the OMPs due to change in pH [5]. Also, compared to the percent-removal of OMPs, the rate constants can be readily coupled with an ideal water reactor (plug-flow, continuous stirred tank) to predict process performance.

Therefore, in this chapter, QSAR models based on ozone/hydroxyl radical rate constants (k_{O3}, k_{OH}) were developed. A sequential statistical approach was followed: correlation analysis, principal component analysis and finally a multi-linear regression equation with ozonation-relevant descriptors. Later, the models were also validated. Correlation analysis helps to identify the significant molecular descriptors; principal component analysis helps to study the data/cluster patterns and reduce redundant variables; and multi-linear regression creates a QSAR model in the form of an equation between the rate constants and the descriptors. Finally, a validation is performed to testify that the QSAR model can be used to make predictions. Also, based on the substituent parameters associated with Hammett/Taft constants, i.e., electron-withdrawing groups decrease ozonation, and electron-pumping groups enhance ozonation, in this study, the chemical classification of the OMPs based on their functional groups was carried out.

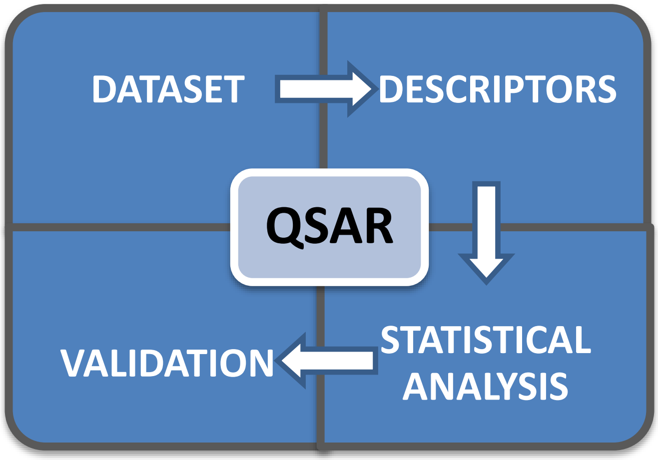


Figure 3.1: Different steps involved in developing a robust QSAR model

3.2. METHODOLOGY

3.2.1. Data Sets

In this study, the k_{OH} values of 83 OMPs and k_{O3} of 40 OMPs from pH 5-8 were chosen for the QSAR modeling. The k_{OH} and k_{O3} values were taken from published scientific articles and the references for all the OMPs are available in *Table 3.1*. Of the 83 OMPs for the k_{OH} QSAR model, 55 were used for the training set and the remaining 28 were used for the test/validation dataset. The k_{OH} values ranged from 0.04 to 18 (10^9) $M^{-1}s^{-1}$. In case of k_{O3} QSAR model, 27 OMPs were used for training set ranging from $5*10^{-4}$ to 10^5 $M^{-1}s^{-1}$ and 13 were used for the test sets. The chemical structures of the OMPs are available in *Appendix A2*

Table 3.1: OMPs and their hydroxyl (k_{OH}) and ozone rate constants (k_{O3})

COMPOUND	**Formula**	**Type**	**$k_{OH}(10^9)$**	**Reference**
1,1,1-Trichloroethane	$C_2H_3Cl_3$	Organic solvent	0.04	[6]
Dibromomethane	CH_2Br_2	Organic solvent	0.09	[7]
1,1,2-Trichloroethane	$C_2H_3Cl_3$	Organic solvent	0.11	[6]
Bromoform	$CHBr_3$	Sedative	0.11	[7]
1,1-Dichloroethylene	$C_2H_2Cl_2$	Polymerisation	0.13	[6]
1,2-Dichloroethane	$C_2H_4Cl_2$	Polymerization	0.2	[6]
1,2-Dibromoethane	$C_2H_4Br_2$	Pesticide	0.26	[6]
Butanol-tert	$C_4H_{10}O$	Gasoline	0.48	[8]
Methanol	CH_3OH	Organic solvent	0.83	[8]
Propan-2-ol	C_3H_7OH	Disinfectant	1.6	[8]
Tetrachloroethylene	C_2Cl_4	Organic solvent	2	[9]
Butan-1,3-diol	$C_4H_{10}O_2$	Organic solvent	2.2	[10]
Ethanol	C_2H_5OH	Organic solvent	2.2	[8]
Dioxane	$C_4H_5O_2$	Organic solvent	2.5	[11]
Propanol	C_3H_7OH	Organic solvent	2.5	[10]
Trichloroethylene	C_2HCl_3	Anesthetic	2.9	[9]
dichloroethylene-cis	$C_2H_2Cl_2$	Organic solvent	3.8	[9]
Diethylether	$C_4H_{10}O$	Organic solvent	4.2	[11]
Butanol	C_4H_9OH	Fuel	4.3	[10]
Benzene Sulfonate	$C_6H_5O_3S$-	Pharmaceutical	4.7	[12]
Imidazole	$C_3H_4N_2$	Pharmaceutical	5.2	[13]
1,4-Dichlorobenzene	$C_6H_4Cl_2$	Pesticide	5.4	[14]
Anisole	$CH_3OC_6H_5$	Pharmaceutical	5.4	[15]
Fumaric acid	$C_4H_4O_4$	Pharmaceutical	6	[16]

Maleic acid	$C_4H_4O_4$	Pharmaceutical	6	[16]
Styrene	C_8H_8	Polymerization	6	[17]
Benzoate ion	$C_7H_5O_2$	Analgesic	6.2	[18]
Ranitidine	$C_{13}H_{22}N_4O_3S$	Antibiotic	6.2	[13]
Pyridoxine	$C_8H_{11}NO_3$	Pharmaceutical	6.3	[19]
tert-Butylhydroquinone	$C_{10}H_{14}O_2$	Anti-oxidant	6.3	[20]
1,3,5-Trimethylbenzene	C_9H_{12}	Organic solvent	6.4	[21]
Acetophenone	C_8H_8O	Pharmaceutical	6.5	[22]
phenol	C_6H_5OH	Cosmetics	6.6	[23]
1,2,3-Trimethylbenzene	C_9H_{12}	Organic solvent	7	[21]
xylene-p	C_8H_{10}	Polymerization	7	[21]
Azauridine	$C_8H_{11}N_3O_6$	Pharmaceutical	7.2	[24]
Chloroquine	$C_{18}H_{26}ClN_3$	Pharmaceutical	7.3	[25]
Ethyl benzene	C_8H_{10}	Organic solvent	7.5	[21]
propylbenzene(iso)	C_9H_{12}	Organic solvent	7.5	[21]
xylene-m	C_8H_{10}	Organic solvent	7.5	[21]
Benzotriazole	$C_6H_5N_3$	Pharmaceutical	7.6	[26]
Carnosine	$C_9H_{14}N_4O_3$	Pharmaceutical	9	[27]
4-chlorophenol	C_6H_5ClO	Pharmaceutical	9.3	[28]
Nitrofurantoin	$C_8H_6N_4O_5$	Pharmaceutical	9.3	[29]
Feruloylputrescine	$C_{14}H_{20}N_2O_3$	Pharmaceutical	10	[30]
Nifuraldezone	$C_7H_6N_4O_5$	Pharmaceutical	10	[29]
Nifuroxime	$C_5H_4N_2O_4$	Pharmaceutical	10	[29]
Tiotidine	$C_{10}H_{16}N_8S_2$	Pharmaceutical	10	[13]
Dipyridamole	$C_{24}H_{40}N_8O_4$	Pharmaceutical	12	[31]

Primaquine	$C_{15}H_{21}N_3O$	Pharmaceutical	12	[25]
Riboflavin	$C_{17}H_{20}N_4O_6$	Pharmaceutical	12	[32]
Quercetin	$C_{15}H_{10}O_7$	Pharmaceutical	15	[33]
Epigallocatechingallate(EGCG)	$C_{22}H_{18}O_{11}$	Pharmaceutical	17	[33]
Piroxicam	$C_{15}H_{13}N_3O_4S$	Pharmaceutical	17	[34]
Silybin	$C_{25}H_{22}O_{10}$	Pharmaceutical	18	[35]
Range			**17.96**	
Mean			**6.23**	
Ozone (k_{O3}) rate constants				
Compound	**Formula**	**Type**	**$k_{O3}(M^{-1}s^{-1})$**	**Reference**
Dalapon	$C_3H_4Cl_2O_2$	Pesticide	0.0005	[36]
Propionic acid	C_2H_5COOH	Pesticide	0.002	[37]
Adipic acid monoethyl ester	$C_7H_{12}O_4$	Polymerisation	0.003	[36]
Succinic acid	$C_4H_6O_4$	Pharmaceutical	0.003	[37]
Tert-Butanol	$C_4H_{10}O$	Gasoline	0.003	[38]
Methanol	CH_3OH	Organic solvent	0.024	[38]
Cyanic acid	$CHNO$	Organic synthesis	0.01	[39]
Glyoxylic acid	$C_2H_2O_3$	Pharmaceutical	1.9	[37]
Propan-2-ol	C_3H_7OH	Disinfectant	1.9	[38]
Endothall	$C_8H_{10}O_5$	Pesticide	2	[36]
Pyridine	C_5H_5N	Pharmaceutical	3	[37]
Octanal	$C_8H_{16}O$	Perfumes	8	[38]
Fenoprop	$C_9H_7Cl_3O_3$	Pesticide	8.9	[36]
Nitrosodimethylamine	$C_2H_6N_2O$	Fuel	10	[37]
2-isopropyl-3-methoxypyrazine	$C_8H_{12}N_2O$	Odorant	50.2	[40]

Picloram	$C_6H_3Cl_3N_2O_2$	Pesticide	100	[36]
Methylamine	CH_5N	Organic solvent	280	[37]
Beta-Alanine	$C_3H_7NO_2$	Pharmaceutical	310	[37]
1-phenoxy-2-propanol	$C_9H_{12}O_2$	Pharmaceutical	320	[41]
Oxamyl	$C_7H_{13}N_3O_3S$	Pharmaceutical	620	[36]
Alpha-Alanine	$C_3H_7NO_2$	Pharmaceutical	640	[37]
Atenolol	$C_{14}H_{22}N_2O_3$	Pharmaceutical	1700	[41]
Metoprolol	$C_{15}H_{25}NO_3$	Pharmaceutical	2000	[41]
Beta-Cyclocitral	$C_{10}H_{16}O$	Perfumes	3890	[40]
Dimethylamine	C_2H_7N	Pharmaceutical	20^3	[37]
Trimethylamine	C_3H_9N	Dye agent	50^3	[37]
Propanolol	$C_{16}H_{21}NO_2$	Pharmaceutical	10^5	[41]
Range			**100000**	
Mean			**6664.67**	
Median			**10**	

3.2.2. Molecular Descriptors Computation

All of the molecules were simulated using "ChemBio3D Ultra 12" software. The chemical structures were obtained using the simplified molecular input line entry specification (SMILES) notation. "Qikprop, Schrödinger software 2010", "HyperChem 8.0.6", were also used to compute advanced 3-D descriptors which involve structural conformation and orbitals of the molecule such as ionisation potential, electron affinity, energy of molecular orbitals, etc. The semi-empirical method, parameterized model number (PM3) in "HyperChem 8.0.6", was used for molecular structure optimization and computation of molecular orbital energies. The optimizations resulted in low energy, highly stable chemical structures of the OMPs and is a pre-requisite to compute the descriptors such as ionisation potential (IP), electron-affinity (EA), and energy of the molecular orbitals. Several molecular descriptors influencing ozonation/oxidation were chosen for this study and they ranged from simple atom-counts to advanced molecular orbital energies as shown in *Table 2.3* of *Chapter 2*. These descriptors help to understand the ozonation/oxidation mechanism in greater depth.

3.2.3. QSAR model and validation

SPSS version 19.0 was used for the statistical analysis. In the case of the k_{O3}-QSAR model, the k_{O3} values ranged from $5*10^{-4}$ to 10^{5} $M^{-1}s^{-1}$. The k_{O3}-dataset was natural logarithm (ln) transformed in order to reduce the range of values and symmetrize the responses [42-44]. As stated earlier, the development of the QSAR model took a sequential approach: correlation analysis, principal component analysis, multiple linear regression and finally validation (*Figure 3.2*).

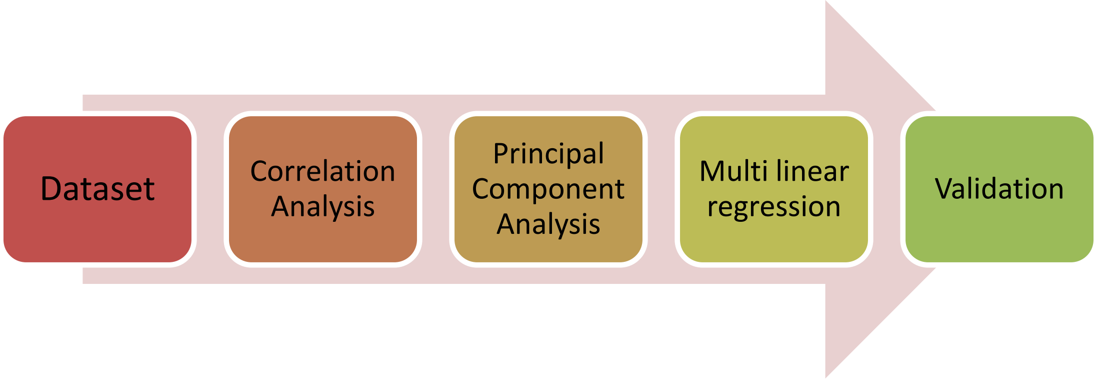


Figure 3.2: Different statistical steps involved in developing the QSAR model

The correlation analysis was performed to analyze the correlation between the descriptors and the rate constants. Next, a principal component analysis (PCA) was performed for variable reduction, intercorrelation and data interpretation. In PCA, variables which describe the same property cluster together, hence it is easy to describe the predicted property with a lesser number of variables [45, 46]. Lastly, multiple linear regression (MLR) was used as the final step in the model building process. The model involves an explicit function relating the set of predictors to the dependent variable [42, 47, 48]. The performance-indices such as r^2, F-statistic, and p-value were analyzed. r^2 describes the proportion of variation in the dependent variable (rate-constants) explained by the regression model, ranging from 0 to 1. Values above 0.5 indicate that the model fits the data well and the model explains a large proportion of the variability in the response. The F-statistic and p-values are used to reject the null-hypothesis. The null-hypothesis proposes that no statistical significance exists in the relationship between the dependent and independent variables. The F-statistic is the ratio of regression and residual mean squares. Residuals are the difference between the observed and predicted values and are computed using the formula:

$$\text{F-statistic} = \frac{\Sigma\ (\hat{y} - \bar{y})^2}{\Sigma\ (y - \hat{y})^2} = \frac{\text{Regression mean square}}{\text{Residual mean square}} \tag{1}$$

y: the dependent variable (rate-constant); $\hat{y}$: predicted value; $\bar{y}$: sample mean

The p-value/significance is the conditional probability that a relationship as strong as the one observed in the data would be present, if the null-hypothesis is true. Therefore, a low p-value (< 0.05) is required to build a statistically significant MLR-based model. In other words, a low p-value (< 0.05) indicates a 95% confidence in the linear relationship between the dependent and independent variable(s) involved in the regression analysis.

Validations of the QSAR models were performed using MOBYDIGS software. The two different types of validation: internal and external were performed on both the models. The theory associated with the internal and external validations are provided in *Chapter 2* (*section.2.2.3*).

In the case of external validation for the QSAR models, a new dataset, not involved in the model-building process, was used. The external data was representative of the same chemical domain as the training set with parameters such as the pH (5-8), and range of k_{OH} values (0.04 ~18* (10^9) $M^{-1}s^{-1}$), k_{O3} values ($5*10^{-4}$ to 10^5 $M^{-1}s^{-1}$) involved in the QSAR model datasets. The external validation (Q^2_{ext}) for the model was computed [49] using the formula:

$$Q^2_{ext} = 1 - \frac{\sum_{i=1}^{valid} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{valid} (y_i - \bar{y}_{tr})^2} \qquad (3)$$

y_i and $\hat{y}_i$ are the measured and predicted values of the dependent variable (over the validation set), and $\bar{y}_{tr}$ is the averaged value of the dependent variable for the training set; the summations cover all of the compounds in the validation set.

The residual plots were also studied as a measure of the validation test. Residuals are the difference between the experiment observed value and the QSAR-model predicted value. If the points in the residual plots are scattered/random without any pattern, it indicates that the regression model is appropriate for the data [50, 51].

3.3. RESULTS AND DISCUSSION

3.3.1. Correlation Analysis

The correlations between the predicted response (k_{OH} or k_{O3}) and the various descriptors were analyzed to check the strength of relationship between the rate-constants (k_{OH}, k_{O3}) and the individual descriptors, and to identify potential collinearity among the descriptors themselves. There were 20 out of 35 descriptors which exhibited correlation coefficients (r) above 0.5 for the k_{OH} dataset (*see Appendix B2*). Molecular properties related to atom-counts such as double-bond equivalence (DBE), number of carbon-carbon double bonds (#C=C), number of aromatic rings, number of ring atoms, number of halogens (#X), number of alcoholic groups (#OH), number of non-hydrogen atoms (#nonHatoms), number of rotational bonds (#rotr), number of metabolites (#metab), number of donor hydrogen bonds (#donorHB), and acceptor hydrogen bonds (#accptHB) correlated well with the rate-constants. Molecular surface area based descriptors such as solvent accessible surface area (SASA), pi-component of SASA (PISA), weakly-polar component of SASA (WPSA), hydrophilic component of SASA (FISA), polar surface area (PSA), polarizability (QPpolrz), ionisation potential (IP), and dipole moment correlated well. In the case of k_{O3} dataset, only 3 out of 35 descriptors showed correlation coefficients (r) above 0.5 with $\ln k_{O3}$. Energy difference between lowest unoccupied and highest occupied molecular orbitals ($E_{LUMO} - E_{HOMO}$), oxygen-carbon ratio (O/C), and ionisation potential (IP) correlated well with the rate constants.

3.3.2. Principal Component Analysis

A principal component analysis (PCA) was performed to further reduce the variables and choose the most significant variables (*Figure 3.3*) [45]. The PCA analysis resulted in final three components that help to explain 82% of variance within the k_{OH} dataset and

75% within the k_{O3} dataset (*Figure 3.3,* pie-charts). Variables with correlation co-efficients (r) greater than 0.5 with the rate-constants (k_{OH}) were chosen for the PCA analysis. Upon PCA analysis for the k_{OH}-dataset, component 1 was dominated by properties (red circles in the loading plot and blue color in the pie-chart) such as polar surface area, donor/acceptor hydrogen bonds, dipole moment, #rotational bonds etc. Component 2 was dominated and had clusters (blue circles in loading plots and brown color in pie chart) describing the double-bond nature of the OMPs (DBE, #C=C, #arom, PISA), ionisation potential (IP), #aromatics and component 3 described the halogens in the compounds (#X, WPSA) (red circles in loading plot and green color of pie chart) (*Figure. 3.3(a, b)*). In the case of the k_{O3}-dataset, since only ionisation potential (IP), oxygen to carbon ratio (O/C) and energy difference between lowest unoccupied and highest occupied molecular orbitals had correlation coefficients (r) greater than 0.5 with $\ln k_{O3}$, the descriptors used for the k_{OH}-dataset PCA-analysis were also considered for the PCA in order to study the data pattern. Component-1 was dominated by clusters describing the double bond nature of OMPs (DBE, PISA, #C=C, #aromatic rings), component 2 was dominated by polar surface area and hydrophilic surface area (FISA), and component 3 was dominated by clusters describing the halogens in the OMPs (weakly polar component of solvent accesible surface area (WPSA) and number of halogens (#X) (*see Figure 3.3(c, d*)).

3.3.3. Multiple-Linear Regression (MLR) based QSAR models

The MLR-based standardized QSAR model for the k_{OH}-and k_{O3} datasets (*Table 3.2*) were defined by descriptors: double-bond equivalence (DBE) and weakly-polar component of solvent accessible surface area (WPSA) and ionisation potential (IP) (*Figure 3.4*).

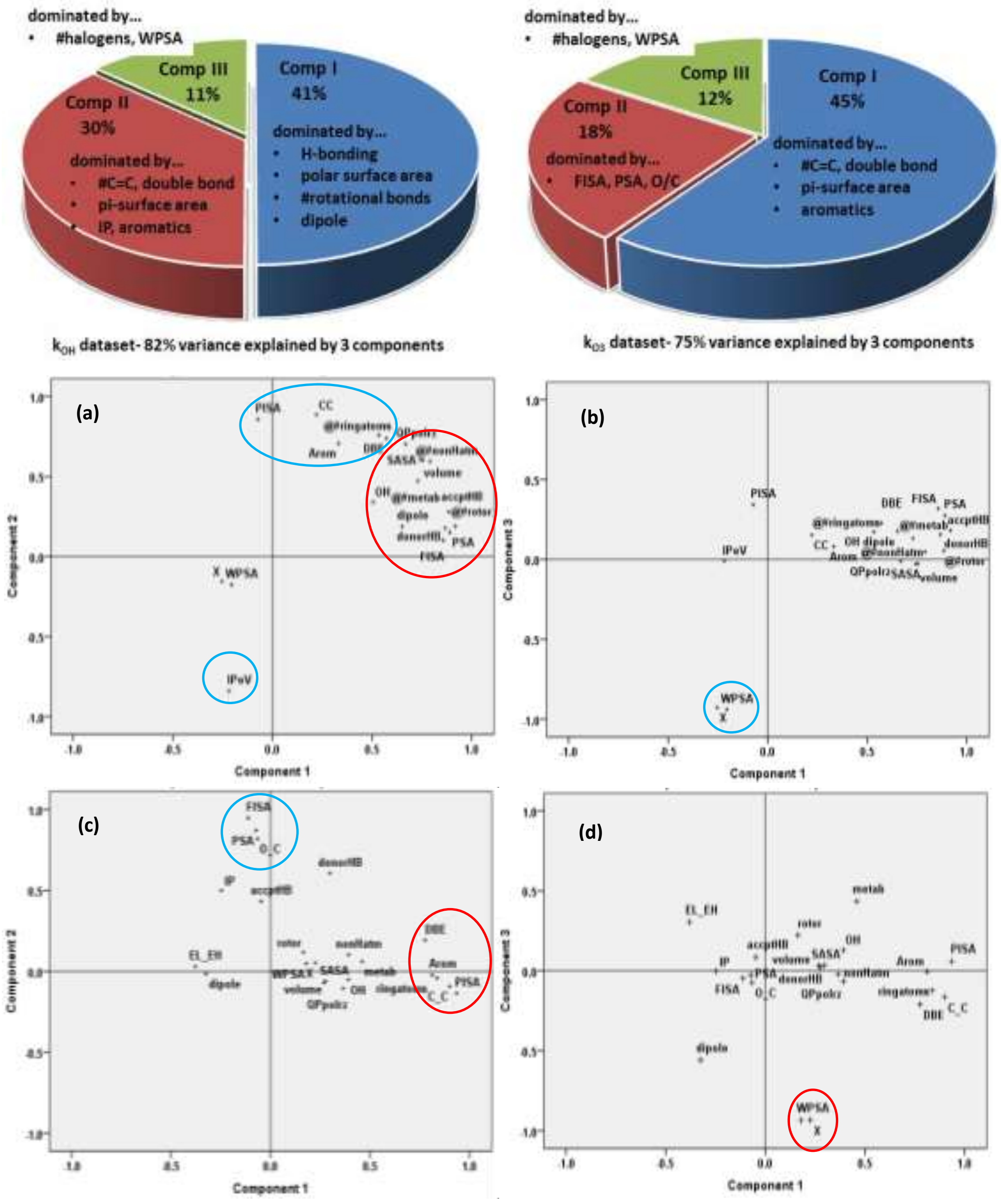


Figure 3.3: Loading plots for the k_{OH} and k_{O3} dataset

(3.3a, 3.3b for k_{OH} dataset; 3.3c, 3.3d for k_{O3} dataset) Pie-chart distribution of the dominant properties of the 3 components for k_{OH} and k_{O3} dataset

The hydroxyl radical rate-constant (k_{OH})/ozone rate constant (k_{O3}) were the dependent variables. In the case of ozone rate constants (k_{O3}), the k_{O3} was natural logarithm 'ln' transformed and the $\ln k_{O3}$ values ranged from -7.6 to 11.51 $M^{-1}s^{-1}$

The natural logarithm (ln) helped to symmetrize the responses and obtain a better model-fit response (r^2). Double bond equivalence (DBE) helps to determine the number of rings, double bonds or triple bonds present in a compound. For systems which contain carbon, hydrogen, halogens, and nitrogen, the DBE is calculated as:

$$\text{DBE} = \#\text{ of carbons} - \frac{1}{2}(\#\text{hydrogens} + \#\text{ halogens} - \#\text{nitrogens}) + 1$$

For a ring system or a double-bond, DBE equals 1. For a triple-bond, DBE equals 2. The OMPs which contain electron-donating carbon systems undergo ozonation readily since the electrons can be readily donated to ozone, an electrophile [1, 52]. The olefins with high electron-density (high DBE) undergo the 'Criegee mechanism' wherein upon ozone reaction, a ozonide is formed, followed by a carbonyl compound [1]. In the case of activated aromatics (presence of electron-donating groups such as alcohols, alkyls, amines, ethers), the electrophilic reaction takes place and ozone attacks at the ortho/para position resulting in the formation of aromatic alcohols [53]. Hence, olefins/activated aromatics with high DBE increase ozonation efficiency. The weakly polar component of the solvent accessible surface area (WPSA) was computed using "Qikprop, Schrodinger, 2010". WPSA describes compound surface area which comprises halogens. Halogens are electrophilic and withdraw electrons from the chemical-system and thereby make the reactive molecular sites electron-deficient and hence inhibit the ozonation process [52]. Also, due to large size of halogen atoms such as chlorine, bromine, etc., they create a steric hindrance for attack by ozone on the OMPs. Hence, larger WPSA decreases ozonation efficiency. Ionisation potential in electron-volts (eV) is the ease with which an electron can be released from a neutral chemical-system. IP is negatively related to ozonation since ozone is an electrophile and lower IP provides greater electrons for ozone and hence increases ozonation [54, 55].

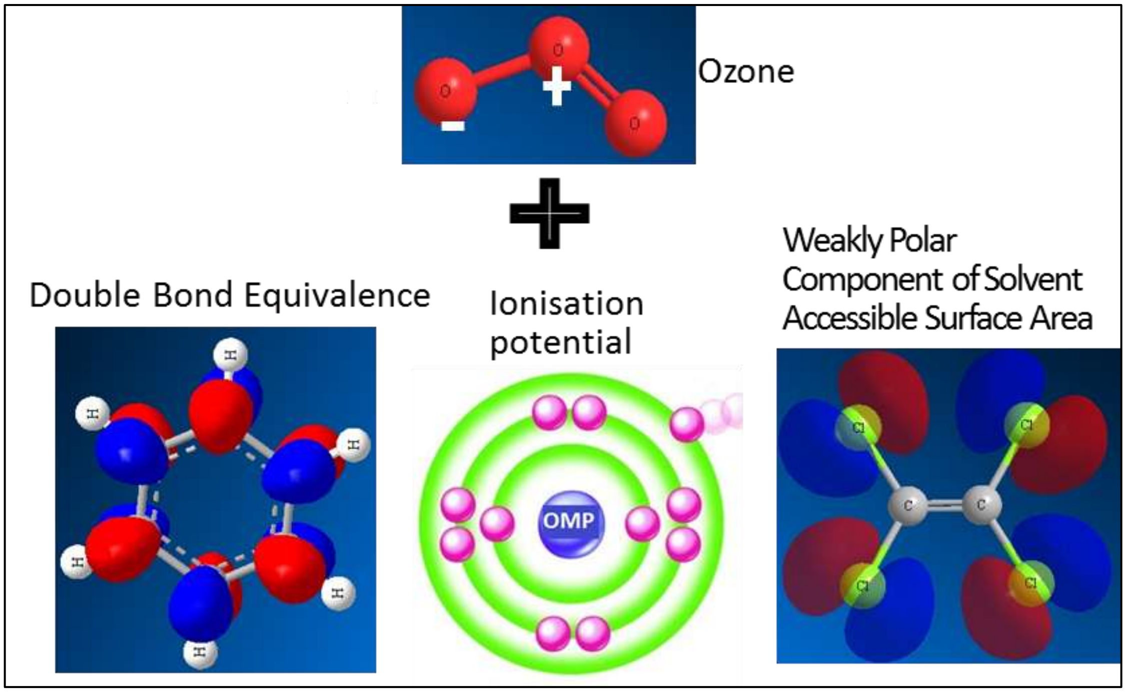


Figure 3.4: Pictorial representation of the QSAR defining molecular descriptors

The statistical criteria for a good model such as coefficient of determination (r^2), F-statistic and p-values were studied. r^2 was 0.918 and 0.832 for k_{OH} and k_{O3} QSAR models, respectively. The F-statistic for k_{OH}-dataset was 290.232, while for the k_{O3} dataset, it was 37.95 (*Table 3.3a, 3.3b*). For both models, the p-values were less than 0.05, indicating more than 95% confidence in the relationship between the descriptors and the rate-constants defining the QSAR models. The high F-statistic value and the low p-value (<0.05) eliminate the null-hypothesis. The MLR-output for both k_{O3} and k_{OH} data are shown in *Tables 3.3a and 3.3b*. In order to avoid multi-collinearity, the inter-correlations among the QSAR-descriptors were studied for both the models. The collinearity between descriptors should be as low as possible, but must be significantly lower than the statistical fit (r^2) of the QSAR model [3]. In this study, descriptors with intercorrelations lower than 0.5 were considered in the QSAR model. The correlation coefficient, in the case of the k_{OH}-QSAR model between DBE and WPSA, was only -0.416, hence, there was no inter-correlation. In the case of the k_{O3}-QSAR model, inter-correlations between DBE and WPSA (0.316), DBE and IP (-0.264), and WPSA and IP (-0.143) were low, thus eliminating inter-correlations between the descriptors.

Table 3.2: Standardized QSAR model equations for rate constants (k_{O3}, k_{OH})

AOP (k_{OH} - data)	$k_{OH} = 2.153 + 0.895(DBE) – 0.134(WPSA)$ $r^2 = 0.918$; $Q^2_{LOO} = 0.909$; $Q^2_{BOOT} = 0.903$
O_3 (k_{O3}- data)	$\ln k_{O3} = 43.765 + 0.195(DBE) - 0.28(WPSA) – 0.855(IP)$ $r^2 = 0.832$; $Q^2_{LOO} = 0.78$; $Q^2_{BOOT} = 0.72$

Units of WPSA: square Angstroms (A^2); IP: electron-volts (eV)

The prediction plots for the k_{OH} and k_{O3} QSAR models are shown in *Figure 3.5*. The k_{OH} model (*Figure 3.5a*) has an even spread of the data comprising OMPs from low to high k_{OH} values. In the k_{O3} plot (*Figure 3.5b*), the $\ln k_{O3}$ exhibited both positive and negative values and an even spread of the data throughout the plot. The dotted line is the ideal reference line (45^0, r^2 =1) and the bold line is the model-regression line. The reference (dotted) line is used to check the degree of over- or under-prediction observed by the MLR model.

Table 3.3a: MLR output for k_{OH} data QSAR model

Model Summary

Model	r	r^2	Adjusted r^2	Std. Error of the Estimate
1	.958[a]	.918	.915	1.3144130

Predictors: (Constant), WPSA, DBE

Dependent Variable: $k_{OH}(10^9)$

ANOVA

Model		Sum of Squares	df	Mean Square	F	Significance
1	Regression	1002.856	2	501.428	290.232	.000[a]
	Residual	89.839	52	1.728		
	Total	1092.695	54			

Predictors: (Constant), WPSA, DBE

Dependent Variable: $k_{OH}(10^9)$

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Significance
		B	Std. Error	Beta		
1	(Constant)	2.153	.329		6.538	.000
	DBE	1.102	.054	.895	20.463	.000
	WPSA	-.008	.003	-.134	-3.053	.004

Dependent Variable: $k_{OH}(10^9)$

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.333970	18.681522	6.233636	4.3094556	55
Residual	-2.4031720	3.3072865	.0000000	1.2898424	55
Std. Predicted Value	-1.369	2.889	.000	1.000	55
Std. Residual	-1.828	2.516	.000	.981	55

Dependent Variable: $k_{OH}(10^9)$

Table 3.3b: MLR output for the k_{O3} data QSAR model

Model Summary

Model	r	r^2	Adjusted r^2	Std. Error of the Estimate
1	.912[a]	.832	.810	2.44

Predictors: (Constant), IP(eV), WPSA, DBE

Dependent Variable: $\ln k_{O3}$

ANOVA

Model		Sum of Squares	df	Mean Square	F	Significance
1	Regression	679.598	3	226.533	37.959	.000[a]
	Residual	137.258	23	5.968		
	Total	816.856	26			

Predictors: (Constant), IP(eV), WPSA, DBE

Dependent Variable: $\ln k_{O3}$

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Significance
		B	Std. Error	Beta		
1	(Constant)	43.765	4.575		9.565	.000
	DBE	.554	.263	.195	2.109	.041
	WPSA	-.030	.010	-.280	-3.096	.005
	IP(eV)	-4.188	.435	-.855	-9.624	.000

Dependent Variable: $\ln k_{O3}$

Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-5.194473	12.012854	2.461556	5.1125713	27
Residual	-3.5548961	3.9351659	.0000000	2.2976439	27
Std. Predicted Value	-1.497	1.868	.000	1.000	27
Std. Residual	-1.455	1.611	.000	.941	27

Dependent Variable: $\ln k_{O3}$

The data-points which scatter far away from the reference line are either over- or under-predicted whereas the data-points in close alignment to the reference line are well-predicted. In the case of both k_{OH} and k_{O3} models, the reference and regression lines were very close to each other, indicating a low degree of over- or under-prediction.

Figure 3.5a and 3.5b also represent classifications of OMPs based on their structure and functional groups, and their influence on ozonation. The prediction plot for the k_{OH}-dataset (*Figure 3.5a*) is dominated by haloalkanes (trichloroethanes, dibromomethane, (tetra-, tri-, dichloro-), ethenes etc.), and ethers (diethyl ether, dioxane) in the lower end of the scatter plot. These classes of OMPs are least reactive to ozonation since the halogens/ethers present on the OMPs withdraw the electrons and make the system electron-deficient. The middle section is dominated by aromatics and alkenes (styrene, pyridoxine, tetra-butyl hydroquinone, etc.). The higher end is dominated by more activated aromatics (azauridine, silybin, piroxicam) containing functional groups such as primary and secondary amines ($-NH_2$, -NH-) and alcohols (-OH) which activate (readily donate electrons to) the benzene molecule and hence easily undergo ozonation.

In case of the k_{O3} dataset (*Figure 3.5b*), the lower end is dominated by haloacids and esters (dalapon, adipic acid monoethyl ester, etc.), the middle section is governed by nitro ($-NO_2$) (nitrosodimethylamine), cyano (-CN) (cyanic acid), aliphatic aldehydes (octanal), and deactivated (presence of halogens (any electron withdrawing groups) on benzene) aromatics (picloram, 2-isopropyl-3-methoxypyrazine, etc.). The upper section consists of activated aromatics, i.e., the presence of primary (1^o) and secondary (2^o) amines, alcohols ((alpha-, beta-), and alanines, (di-, tri-) methylamines, propranolol).

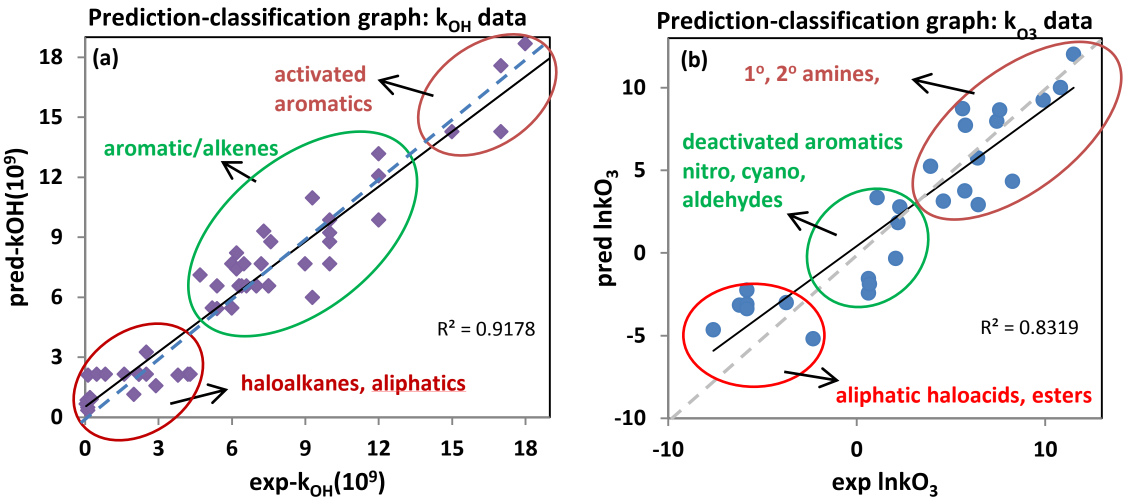


Figure 3.5: QSAR prediction and chemical classification plots for k_{OH} and k_{O3} data
Prediction plots (a), (b): QSAR model values vs experimental values for k_{OH} and k_{O3} data. The dotted line is the reference line ($r^2= 1$) and the bold line is the model-regression line. The circles represent the QSAR - scatter-plots of OMPs based on their structural attribute

3.3.4. Validation

The internal validations performed using "MOBYDIGS" software consisted of leave-one-out (LOO) and bootstrap methods. The Q^2_{LOO} for the k_{OH} dataset was 0.909 and the Q^2_{BOOT} was 0.903. In the case of the k_{O3}-dataset, Q^2_{LOO} was 0.78 and Q^2_{BOOT} was 0.72. In all of the cases, the validation performance indices, Q^2 values, were greater than 0.5, which indicated that the models had good predictive powers.

External validations were performed for both k_{OH} and k_{O3} datasets. The compounds along with their rate constants (k_{OH}), compound properties, and the parameters required to compute the Q^2_{ext} according to equation 3, i.e., *difference between the measured and predicted values of the dependent variable* (y_i-$\hat{y}_i$) and *difference between the measured and average value of the training set* (y_i-$\bar{y}_{tr}$) *used for the external validation*; are listed in *Table 3.4.*

Table 3.4: External validation dataset for k_{OH} and k_{O3} dataset: QSAR descriptors

COMPOUND	k_{OH}(*10^9)	predicted	DBE	WPSA	$(y_i-\hat{y}_i)^2$	$(y_i-\bar{y}_{tr})^2$	Reference
Taurine	0.038	1.77	0	0.742	3.00	38.38	[30]
Chloroform	0.05	0.38	0	213.179	0.11	38.23	[7]
Trichloroacetic acid	0.06	1.71	1	182.926	2.71	38.11	[56]
Spermidine	0.12	1.77	0	0	2.74	37.37	[30]
Formaldehyde	0.76	2.99	1	0	4.96	29.95	[57]
Butan-2-one	0.9	2.99	1	0	4.35	28.44	[10]
Propylacetate	1.4	2.99	1	0	2.52	23.36	[10]
MTBE	1.9	1.85	0	0	0.00	18.77	[58]
Propanal	2.2	2.99	1	0	0.62	16.27	[59]
sucrose	2.3	4.13	2	0	3.34	15.47	[60]
2-Methylisoborneol	3	4.13	2	0	1.27	10.45	[61]
Penicillamine	3.3	2.61	1	53.015	0.47	8.60	[62]
Clofibric acid a	5	7.04	5	71.454	4.17	1.52	[63]
Toluene	5.1	6.40	4	0	1.70	1.28	[64]
Roxithromycin	5.4	7.54	5	0	4.59	0.69	[65]
Chlorobenzene	6	5.90	4	71.46	0.01	0.05	[66]
Carbofuran	7	8.68	6	0	2.83	0.59	[7]
o-Xylene	7	6.40	4	0	0.36	0.59	[21]
Alachlor	7	7.12	5	60.116	0.01	0.59	[7]
Amikacin	7.2	6.40	4	0	0.64	0.94	[65]
Ibuprofen	7.4	7.54	5	0	0.02	1.36	[67]
Benzene	8	6.40	4	0	2.55	3.12	[66]
Naproxen	10	10.96	8	0	0.92	14.19	[68]
17α-ethinylestradiol	10	12.10	9	0	4.40	14.19	[67]
Naphthalene	12	9.82	7	0	4.75	33.26	[69]
Carmine	13	16.65	13	0	13.35	45.79	[70]

Benzoflavine	14	17.79	14	0	14.39	60.33	[71]
Quinidine	18	13.24	10	0	22.69	138.46	[72]
Range	**17.96**	**17.41**	∑		**103.44**	**620.35**	
Mean	**5.65**	**6.43**					

External Validation k_{O3} dataset

COMPOUND	k_{O3}	$\ln k_{O3}$	predicted	DBE	WPSA	IP	$(y_i-\hat{y}_i)^2$	$(y_i-\bar{y}_{tr})^2$	Ref.
Glutaric acid	0.008	-4.828	-3.423	2	0	11.53	1.97	51.959	[37]
methanol	0.024	-3.729	-5.03	0	0	11.65	1.67	37.328	[38]
Lindane	0.04	-3.218	-5.848	1	325.12	9.65	6.92	46.007	[36]
MTBE	0.14	-1.966	1.382	0	0	10.12	11.21	18.889	[58]
p-chlorobenzoic acid	0.15	-1.897	1.048	5	71.37	10.35	8.67	18.293	[36]
ethanol	0.51	-0.67	-2.0936	0	0	10.95	2.02	17.955	[36]
creatinine	2	0.693	2.081	3	0	10.35	1.93	2.130	[37]
butylamine	340	5.828	4.235	0	0	9.44	2.54	5.129	[37]
napthalene	3800	8.243	9.534	5	0	8.84	1.67	21.891	[2]
glyphosate	8200	9.012	7.604	0.5	1.47	8.69	1.98	29.679	[36]
flouranthene	9500	9.159	13.743	12	0	8.76	21.01	31.305	[2]
phenanthrene	10000	9.21	7.509	10	0	9.98	2.89	31.88	[2]
anthracene	27000	10.204	12.66	10	0	8.75	6.03	44.08	[2]
Range	**26999.99**	**15.03**	**19.59**			∑	**87.81**	**352.22**	
Mean	**4526.37**	**2.77**	**3.34**						

The external validations for k_{OH} and k_{O3} datasets resulted in a Q^2_{ext} value of 0.83 and 0.75, respectively. The k_{OH} and k_{O3} dataset prediction plots for external data are shown in *Figures 3.6a and 3.6b*. The dotted line is the ideal-reference line (45^0, $r^2=1$) which is used to indicate the degree of over- or under- prediction by the model regression line. *Figure 3.6a and 3.6b* also show the classification of the OMPs based on their structures. In the case of the k_{OH} dataset (*Figure 3.6a*) the lower section consists of haloalkanes (chloroform, trichloroacetic acid, etc.), ethers (MTBE), and carbonyls (formaldehyde, propanal, buatan-2-one, etc.); the middle section consists of acids (clofibric acid) and several aromatics (ibuprofen, benzene, toluene etc.); and finally the upper section consists of activated (mainly conjugated, amines, triple bonds) aromatics (carmine, benzoflavine quinidine, 17α-ethinyl estradiol). In case of the k_{O3} dataset (*Figure 3.6b*) the lower section consists of haloacids and ethers, and the upper section consists of aromatics and alkenes. The influence of these functional groups has been discussed earlier in relation to the classification for OMPs of the k_{OH} and k_{O3} datasets (*section 3.3.3*). The internal and external validations performed on the k_{OH} and k_{O3} datasets and the associated Q^2 values for leave-one-out, bootstrap and the external validations are shown in *Table 3.5*.

Table 3.5: Internal and external validation of the k_{OH} and k_{O3} QSAR models

DATASET	INTERNAL	EXTERNAL
k_{OH} dataset	Leave-one-out: $Q^2_{LOO} = 0.909$ Bootstrap: $Q^2_{boot} = 0.903$	$Q^2_{ext} = 0.856$
k_{O3} dataset	Leave-one-out: $Q^2_{LOO} = 0.78$ Bootstrap: $Q^2_{boot} = 0.72$	$Q^2_{ext} = 0.823$

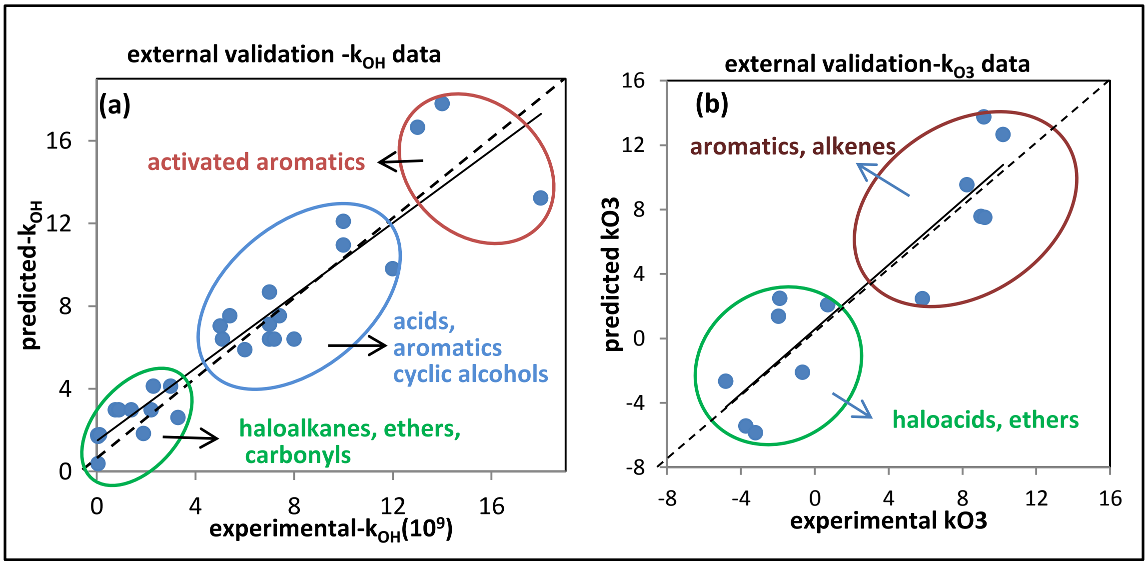


Figure 3.6: External validation plots for k_{OH} and k_{O3} data
(a) External validation of the kOH- QSAR models and the OMPs classification
(b) External validation for the kO3-QSAR models and the OMPs classification

The residual plots for the k_{OH} and k_{O3} data were plotted (*Figure 3.7*). Residuals are the difference between the observed and the predicted values. A random scatter plot was observed in both of the cases which indicated that the regression model was appropriate for the data.

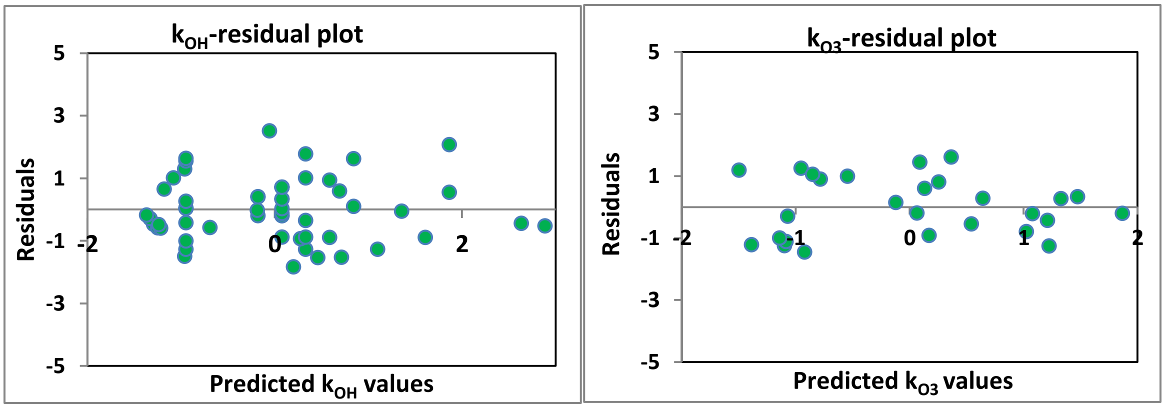


Figure 3.7: Residual plots of k_{OH} and k_{O3} QSAR models

3.4. CONCLUSIONS

QSAR models for ozonation/oxidation, predicting rate-constants (k_{OH} and k_{O3}), were developed. Among the several molecular descriptors which influence (advanced) oxidation, DBE and WPSA chemically and statistically defined the k_{OH}-QSAR model. For the k_{O3}-QSAR model, three descriptors, DBE, WPSA and IP, defined the model. DBE focuses on the double-bond nature of the OMPs which enhances ozonation efficiency. WPSA focuses on the surface area occupied by halogens, and IP represents the energy required to remove an electron from a neutral atom. An increase in both WPSA and IP decreases ozonation efficiency. The chemical and statistical significance of the descriptors were studied. In the case of the k_{OH}-QSAR model, DBE influenced the model more than WPSA. For the k_{O3}-QSAR model, IP was more influential than WPSA and DBE. Statistical criteria such as the co-efficient of determination, r^2, p-values, and F-statistic were also evaluated, and the models yielded good results. The internal validations (leave-one-out and bootstrap) for the k_{OH} and k_{O3} models yielded high Q^2 values (>0.5), indicating the robustness of the models. External validations were performed for both the QSAR models, resulting in a good Q^2_{EXT} (>0.5). The validation studies helped to conclude that the models were robust and predictive. The residual plots were also studied and random scatter plots were obtained for both the k_{OH} and k_{O3} QSAR models.

The applicability domains for k_{OH} - QSAR models are pH (5 to 8) and k_{OH} values (0.04 to 18 (10^9) $M^{-1}s^{-1}$); for the k_{O3} - QSAR model, pH (5 to 8) and k_{O3} values ($5*10^{-4}$ to 10^5 $M^{-1}s^{-1}$).

REFERENCES

1. von Gunten, U., *Ozonation of drinking water: Part I. Oxidation kinetics and product formation.* Water Research, 2003. **37**(7): p. 1443-1467.

2. Trapido, M., Veressinina, Y., Munter, R., *Ozonation and advanced oxidation processes of polycyclic aromatic hydrocarbons in aqueous solutions-a kinetic study*. Environmental technology, 1995. **16**(8): p. 729-740.

3. Cronin, M.T.D.,Schultz, T.W., *Pitfalls in QSAR.* Journal of Molecular Structure: THEOCHEM, 2003. **622**(1): p. 39-51.

4. Lee, Y.,von Gunten, U., *Quantitative structure-activity relationships (QSARs) for the transformation of organic micropollutants during oxidative water treatment.* Water Research, 2012.

5. Sonntag, C.V.C.,von Gunten, U., *Chemistry of Ozone in Water and Wastewater Treatment From Basic Principles to Applications*. 2012: IWA Publishing. 312.

6. Lal, M., Schöneich, C., Mönig, J., Asmus, K.-D., *Rate Constants for the Reactions of Halogenated Organic Radicals.* International Journal of Radiation Biology, 1988. **54**(5): p. 773-785.

7. Haag, W.R.,Yao, C.C.D., *Rate constants for reaction of hydroxyl radicals with several drinking water contaminants.* Environmental Science & Technology, 1992. **26**(5): p. 1005-1013.

8. Motohashi, N.,Saito, Y., *Competitive measurement of rate constants for hydroxyl radical reactions using radiolytic hydroxylation of benzoate.* Chemical and Pharmaceutical Bulletin, 1993. **41**(10): p. 1842-1845.

9. Getoff, N., *Radiation-and photoinduced degradation of pollutants in water. A comparative study.* International Journal of Radiation Applications and Instrumentation. Part C. Radiation Physics and Chemistry, 1991. **37**(5): p. 673-680.

10. Adams, G., Boag, J., Michael, B., *Reactions of the hydroxyl radical. Part 2.—Determination of absolute rate constants.* Transactions of the Faraday Society, 1965. **61**: p. 1417-1424.

11. Thomas, J., *Rates of reaction of the hydroxyl radical.* Transactions of the Faraday Society, 1965. **61**: p. 702-707.

12. Anbar, M.,Hart, E.J., *The Reactivity of Aromatic Compounds toward Hydrated Electrons.* Journal of the American Chemical Society, 1964. **86**(24): p. 5633-5637.

13. Ching, T.L., Haenen, G.R.M.M., Bast, A., *Cimetidine and other H< sub> 2</sub> receptor antagonists as powerful hydroxyl radical scavengers.* Chemico-Biological Interactions, 1993. **86**(2): p. 119-127.

14. Kochany, J.,Bolton, J.R., *Mechanism of photodegradation of aqueous organic pollutants. 2. Measurement of the primary rate constants for reaction of hydroxyl radicals with benzene and some halobenzenes using an EPR spin-trapping method following the photolysis of hydrogen peroxide.* Environmental Science & Technology, 1992. **26**(2): p. 262-265.

15. O'Neill, P., Steenken, S., Schulte-Frohlinde, D., *Formation of radical cations of methoxylated benzenes by reaction with OH radicals, Ti2+, Ag2+, and SO4·-in aqueous solution. An optical and conductometric pulse radiolysis and in situ radiolysis electron spin resonance study.* Journal of Physical Chemistry, 1975. **79**(25): p. 2773-2779.

16. Cabelli, D.,Bielski, B., *A pulse radiolysis study of some dicarboxylic acids of the citric acid cycle. The kinetics and spectral properties of the free radicals formed by reaction with the OH radical.* Zeitschrift für Naturforschung. Teil b, Anorganische Chemie, organische Chemie, 1985. **40**(12): p. 1731-1737.

17. Brede, O., Helmstreit, W., Mehnert, R., *Nanosekunden-Pulsradiolyse von Styrol in wäßriger Lösung.* Journal für Praktische Chemie, 1974. **316**(3): p. 402-414.

18. Wolfenden, B.S.,Willson, R.L., *Radical-cations as reference chromogens in kinetic studies of ono-electron transfer reactions: pulse radiolysis studies of 2, 2'-azinobis-(3-ethylbenzthiazoline-6-sulphonate).* Journal of the Chemical Society, Perkin Transactions 2, 1982(7): p. 805-812.

19. Moorthy, P.,Hayon, E., *One-electron redox reactions of water-soluble vitamins. III. Pyridoxine and pyridoxal phosphate (vitamin B6).* Journal of the American Chemical Society, 1975. **97**(8): p. 2048-2052.

20. Dohrmann, J.K.,Bergmann, B., *Equilibria and rates of redox reactions involving the 2-tert-butyl-1, 4-benzosemiquinone radical in aqueous solution: an investigation by potentiometry, ESR, and pulse radiolysis.* The Journal of Physical Chemistry, 1995. **99**(4): p. 1218-1227.

21. Sehested, K., Corfitzen, H., Christensen, H., Hart, E., *Rates of reaction of oxygen (1-) ions, hydroxyl radicals, and atomic hydrogen with methylated benzenes in aqueous solution. Optical spectra of radicals.* The Journal of Physical Chemistry, 1975. **79**(4): p. 310-315.

22. Neta, P.,Dorfman, L.M., *Pulse radiolysis studies. XIII. Rate constants for the reaction of hydroxyl radicals with aromatic compounds in aqueous solutions.* Advances in Chemical Series, 1968. **81**(15): p. 222-230.

23. Field, R., Raghavan, N., Brummer, J., *A pulse radiolysis investigation of the reactions of bromine dioxide radical (BrO2. cntdot.) with hexacyanoferrate (II), manganese (II), phenoxide ion, and phenol.* The Journal of Physical Chemistry, 1982. **86**(13): p. 2443-2449.

24. Rosenthal, I., Riesz, P., Faraggi, M., *Pulse Radiolysis Studies of Some Aza Analogues of Nucleic Acid Components.* International Journal of Radiation Biology, 1983. **44**(5): p. 463-468.

25. Bors, H., Golenser, J., Chevion, M., Saran, M., *Reductive and oxidative radical reactions of selected anti-malarial drugs.* ln: Oxidative damage and Repair. Chemical, Biological and Medical Aspects 1991: p. 234-240.

26. Naik, D.,Moorthy, P., *Studies on the transient species formed in the pulse radiolysis of benzotriazole.* Radiation Physics and Chemistry, 1995. **46**(3): p. 353-357.

27. Pavlov, A.R., Revina, A.A., Dupin, A.M., Boldyrev, A.A., Yaropolov, A.I., *The mechanism of interaction of carnosine with superoxide radicals in water solutions.* Biochimica et Biophysica Acta (BBA)-General Subjects, 1993. **1157**(2): p. 304-312.

28. Stafford, U., Gray, K.A., Kamat, P.V., *Radiolytic and TiO2-assisted photocatalytic degradation of 4-chlorophenol. A comparative study.* The Journal of Physical Chemistry, 1994. **98**(25): p. 6343-6351.

29. Greenstock, C.,Dunlop, I., *Pulse radiolysis studies of nitrofurans. Radiation chemistry of nifuroxime.* The Journal of Physical Chemistry, 1973. **77**(15): p. 1834-1838.

30. Bors, W., Langebartels, C., Michel, C., Sandermann, H., *Polyamines as radical scavengers and protectants against ozone damage.* Phytochemistry, 1989. **28**(6): p. 1589-1595.

31. Iuliano, L., Pratico, D., Ghiselli, A., Bonavita, M.S., Violi, F., *Reaction of dipyridamole with the hydroxyl radical.* Lipids, 1992. **27**(5): p. 349-353.

32. Kishore, K., Moorthy, P., Guha, S., *Pulse radiolysis study of one electron oxidation of riboflavin.* International Journal of Radiation Applications and Instrumentation. Part C. Radiation Physics and Chemistry, 1991. **38**(1): p. 119-125.

33. Wang, W., Luo, J., Yao, S., Lian, Z., Zhang, J., Lin, N., Fang, R., Hu, T., *Interaction of phenolic antioxidants and hydroxyl radicals.* Radiation Physics and Chemistry, 1993. **42**(4): p. 985-987.

34. Oturan, M.A., Pinson, J., Bizot, J., Deprez, D., Terlain, B., *Reaction of inflammation inhibitors with chemically and electrochemically generated hydroxyl radicals.* Journal of Electroanalytical Chemistry, 1992. **334**(1): p. 103-109.

35. György, I., Antus, S., Földiak, G., *Pulse radiolysis of silybin: one-electron oxidation of the flavonoid at neutral pH.* International Journal of Radiation Applications and Instrumentation. Part C. Radiation Physics and Chemistry, 1992. **39**(1): p. 81-84.

36. Yao, D.C.,Haag, W.R., *Rate constants for direct reactions of ozone with several drinking water contaminants.* Water Research, 1991. **25**(7): p. 761-773.

37. Hoigné, J.,Bader, H., *Rate constants of reactions of ozone with organic and inorganic compounds in water—II: Dissociating organic compounds.* Water Research, 1983. **17**(2): p. 185-194.

38. Hoigné, J.,Bader, H., *Rate constants of reactions of ozone with organic and inorganic compounds in water—I: Non-dissociating organic compounds.* Water Research, 1983. **17**(2): p. 173-183.

39. Hoigné, J., Bader, H., Haag, W., Staehelin, J., *Rate constants of reactions of ozone with organic and inorganic compounds in water—III. Inorganic compounds and radicals.* Water Research, 1985. **19**(8): p. 993-1004.

40. Peter, A.,Von Gunten, U., *Oxidation kinetics of selected taste and odor compounds during ozonation of drinking water.* Environmental Science & Technology, 2007. **41**(2): p. 626-631.

41. Benner, J., Salhi, E., Ternes, T., von Gunten, U., *Ozonation of reverse osmosis concentrate: Kinetics and efficiency of beta blocker oxidation.* Water Research, 2008. **42**(12): p. 3003-3012.

42. Kusic, H., Rasulev, B., Leszczynska, D., Leszczynski, J., Koprivanac, N., *Prediction of rate constants for radical degradation of aromatic pollutants in water matrix: A QSAR study.* Chemosphere, 2009. **75**(8): p. 1128-1134.

43. de Ridder, D.J., Villacorte, L., Verliefde, A.R.D., Verberk, J.Q.J.C., Heijman, S.G.J., Amy, G.L., van Dijk, J.C., *Modeling equilibrium adsorption of organic micropollutants onto activated carbon.* Water Research, 2010. **44**(10): p. 3077-3086.

44. Oberg, T., *A QSAR for the hydroxyl radical reaction rate constant: validation, domain of application, and prediction.* Atmospheric Environment, 2005. **39**(12): p. 2189-2200.

45. Yangali-Quintanilla, V., Sadmani, A., McConville, M., Kennedy, M., Amy, G., *A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes.* Water Research, 2010. **44**(2): p. 373-384.

46. Jolliffe, I., *Principal component analysis.* Encyclopedia of Statistics in Behavioral Science, 2002.

47. Ott, L.,Longnecker, M., *An introduction to statistical methods and data analysis*. 2008: Duxbury Pr.

48. Gramatica, P., *Principles of QSAR models validation: internal and external.* QSAR & Combinatorial Science, 2007. **26**(5): p. 694-701.

49. Gramatica, P., *Evaluation of different statistical approaches for the validation of quantitative structure–activity relationships.* ECVAM, Ispra, 2004. **177**.

50. Fatemi, M., *Prediction of ozone tropospheric degradation rate constant of organic compounds by using artificial neural networks.* Analytica Chimica Acta, 2006. **556**(2): p. 355-363.

51. Wang, Y., Loftin, K., Meyer, M., *Removal of antibiotics from surface and distilled water in conventional water treatment processes.* Journal of Environmental Engineering, 2002. **128**: p. 253.

52. Lei, H.,Snyder, S.A., *3D QSPR models for the removal of trace organic contaminants by ozone and free chlorine.* Water Research, 2007. **41**(18): p. 4051-4060.

53. Langlais, B., Reckhow, D.A., Brink, D.R., *Ozone in water treatment: application and engineering*. 1991: CRC.

54. Güsten, H., *Predicting the abiotic degradability of organic pollutants in the troposphere.* Chemosphere, 1999. **38**(6): p. 1361-1370.

55. Grosjean, D.,Williams, E.L., *Environmental persistence of organic compounds estimated from structure-reactivity and linear free-energy relationships. Unsaturated aliphatics.* Atmospheric Environment. Part A. General Topics, 1992. **26**(8): p. 1395-1405.

56. Maruthamuthu, P., Padmaja, S., Huie, R.E., *Rate constants for some reactions of free radicals with haloacetates in aqueous solution.* International Journal of Chemical Kinetics, 1995. **27**(6): p. 605-612.

57. Anbar, M.,Neta, P., *Tables of bimolecular rate constants of hydrated electrons, hydrogen atoms and hydroxyl radicals with inorganic and organic compounds*. 1964, Israel. ATomic Energy Commission. Soreq Research Establishment, Rehovoth.

58. Acero, J.L., Haderlein, S.B., Schmidt, T.C., Suter, M.J.F., von Gunten, U., *Mtbe oxidation by conventional ozonation and the combination ozone/hydrogen peroxide: Efficiency of the processes and bromate formation.* Environmental Science & Technology, 2001. **35**(21): p. 4252-4259.

59. Mezyk, S.P., *Rate constant and activation energy determination for reaction of e-(aq) and OH with 2-butanone and propanal.* Canadian Journal of Chemistry, 1994. **72**(4): p. 1116-1119.

60. Moore, J., Kemsley, K., Davies, J., *Radiation Biology and Chemistry. Research Developments. ed. by Edwards HE, Navaratnam S, Parson BJ and Phillips GO*. 1979, Elsevier, New York.

61. Zarnoch, J., Ruth, E., Chauncey, W., Schep, R., Glaze, W., *Evaluating Oxidants for the Removal of Model Taste and Odor Compounds from a Municipal Water Supply*. Journal - American Water Works Association 1990. **82**(5).

62. Bolkenius, F.N., Grisar, J.M., De Jong, W., *A Water-Soluble Quaternary Ammonium Analog of α-Tocopherol, that Scavenges Lipoperoxyl, Superoxyl and Hydroxyl Radicals.* Free Radical Research, 1991. **14**(5-6): p. 363-372.

63. Packer, J.L., Werner, J.J., Latch, D.E., McNeill, K., Arnold, W.A., *Photochemical fate of pharmaceuticals in the environment: Naproxen, diclofenac, clofibric acid, and ibuprofen.* Aquatic Sciences-Research Across Boundaries, 2003. **65**(4): p. 342-351.

64. Roder, M., Wojnarovits, L., Földiák, G., *Pulse radiolysis of aqueous solutions of aromatic hydrocarbons in the presence of oxygen.* International Journal of Radiation Applications and Instrumentation. Part C. Radiation Physics and Chemistry, 1990. **36**(2): p. 175-176.

65. Dodd, M.C., Buffle, M.O., Von Gunten, U., *Oxidation of antibacterial molecules by aqueous ozone: moiety-specific reaction kinetics and application to ozone-based wastewater treatment.* Environmental Science & Technology, 2006. **40**(6): p. 1969-1977.

66. Ashton, L., Buxton, G.V., Stuart, C.R., *Temperature dependence of the rate of reaction of OH with some aromatic compounds in aqueous solution. Evidence for the formation of a π-complex intermediate?* Journal of the Chemical Society, Faraday Transactions, 1995. **91**(11): p. 1631-1633.

67. Huber, M.M., Canonica, S., Park, G.Y., Von Gunten, U., *Oxidation of pharmaceuticals during ozonation and advanced oxidation processes.* Environmental Science & Technology, 2003. **37**(5): p. 1016-1024.

68. Huber, M.M., GÖbel, A., Joss, A., Hermann, N., LÖffler, D., McArdell, C.S., Ried, A., Siegrist, H., Ternes, T.A., von Gunten, U., *Oxidation of pharmaceuticals during ozonation of municipal wastewater effluents: a pilot study.* Environmental Science & Technology, 2005. **39**(11): p. 4290-4299.

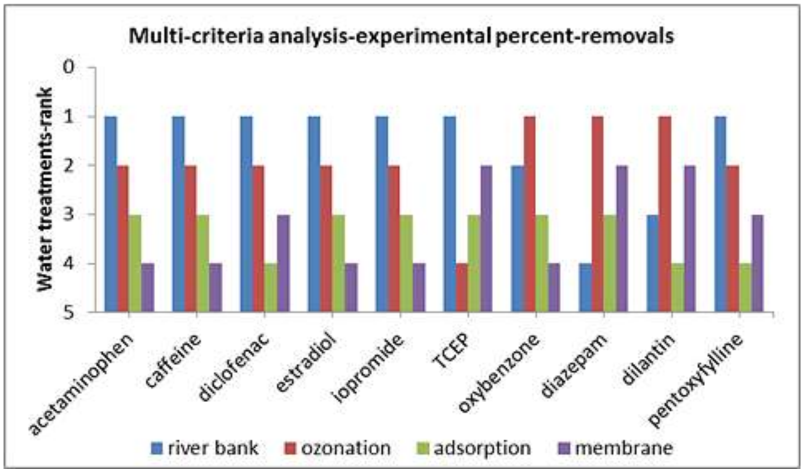
69. Kanodia, S., Madhaven, V., Schuler, R., *Oxidation of naphthalene by radiolytically produced< sup></sup> OH radicals.* International Journal of Radiation Applications and Instrumentation. Part C. Radiation Physics and Chemistry, 1988. **32**(5): p. 661-664.

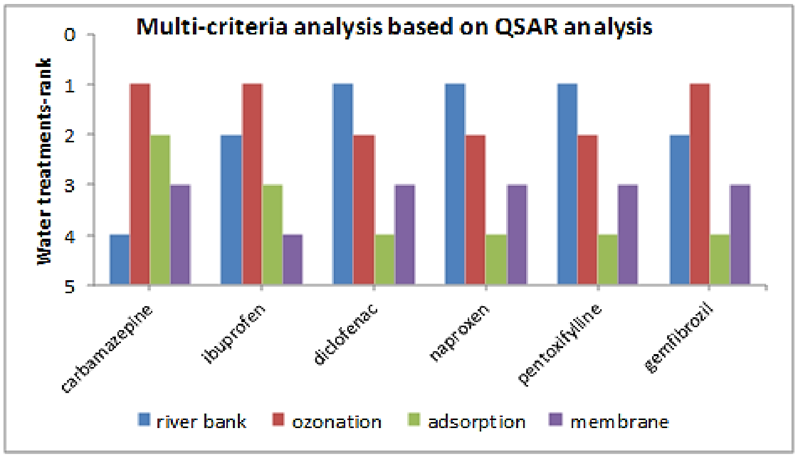
70. Sychev, A.Y., Isak, V., Pfannmeller, U., *Determination of rate constants of hydroxyl radicals with organic and inorganic substances under conditions for the catalytic decomposition of hydrogen peroxide.* Zh. Fiz. Khim, 1979. **53**(11): p. 2790-2793.

71. Buchanan, J., Power, D., Phillips, G., Davies, J.V., *Interaction of Aminoacridines with Nucleic Acids.* International Journal of Radiation Biology, 1978. **33**(6): p. 551-562.

72. Das, K.C.,Misra, H., *Antiarrhythmic agents. Scavengers of hydroxyl radicals and inhibitors of NADPH-dependent lipid peroxidation in bovine lung microsomes.* Journal of Biological Chemistry, 1992. **267**(27): p. 19172-19178.

4. APPROPRIATE DRINKING WATER TREATMENT PROCESSES FOR ORGANIC MICROPOLLUTANTS REMOVAL BASED ON EXPERIMENTAL AND MODEL STUDIES
A MULTI-CRITERIA ANALYSIS





HIGHLIGHTS

- Evaluation of eight water treatment processes for OMPs removal
- MCA under socio-economic, ecological and technical criteria
- River bank filtration and ozonation performed best for OMP removal in the MCA
- RBF is only outranked by ozonation for carbamazepine, diazepam and dilantin
- MCA based on QSAR models, nanofiltration outranked reverse osmosis

This chapter is based on the published paper:

Sairam Sudhakaran, Sabine Lattemann, Gary L. Amy. Appropriate drinking water treatment processes for organic micropollutants removal based on experimental and model studies – A multi – criteria analysis study; *Science of Total Environment*, 442 (2013), 478 – 488

4.1. INTRODUCTION

In the previous chapters, the impact of OMPs on the public health and environment was discussed and an overview provided on regulations which governments/regulatory agencies emphasize in risk assessments associated with OMPs.

Since *Chapter 2* and *Chapter 3* resulted in development of good/robust QSAR models, in this chapter, efforts were focused on developing a decision support system (DSS) based on multi-criteria analysis (MCA) for selection of appropriate water treatment processes for OMPs removal based on QSAR models and experimental studies.

Advanced water treatment processes such as reverse osmosis (RO), nanofiltration (NF), adsorption and oxidation are industry choices for OMPs removal, however, even natural systems such as riverbank filtration (RBF) and constructed wetlands (CW) show good removals of OMPs [1-3]. Usually, decision making involves choosing or judging between several treatment alternatives and ultimately compromising on a result. In order to choose a good and efficient water treatment process, decision making can be facilitated by a DSS based on MCA. MCA has been well-recognized by governmental agencies and regulatory bodies as a standard decision aid procedure [4-8]. MCA has also been used for decision making for pharmaceuticals removal from hospital waste waters [9]. MCA allows the comparison of qualitative (non-monetary, non-metric) criteria, which are often used in environmental contexts, as well as quantitative criteria (e.g. costs). The systematic and transparent approach involved in MCA helps to reproduce and scrutinize the results [8].

The study in this chapter was initiated from the necessity for an approach to decide upon a suitable water treatment process for OMPs removal. Accordingly, the main objective was to construct a DSS for comparing different water treatment processes under various criteria, including percent-removal rates as predicted by QSAR models, and to recommend the best water treatment process. The DSS can provide useful information to set up a treatability experiment for emerging contaminants or support decision making in formulating drinking water treatment trains. However, in practice, there could be more variables affecting the efficiency of contaminant removal and factors related to process selection, which have not been covered in this study.

4.2. METHODOLOGY

MCA analysis was carried out in three stages: MCA-I, MCA-II, and MCA-III. The objective of MCA-I was to study the influence of the various social, economic and technical aspects on the ranking of the different water treatment processes without any influence from the percent-removal rates of these processes, hence, the percent-removal rates have been set to a fixed, arbitrarily chosen value (i.e., 99%). In MCA-II, the percent-removals were selected from pilot-scale experimental studies published in the literature and analyzed in conjunction with the same social, economic and technical aspects as in MCA-I [1]. The MCA-III was built with the predicted percent-removals from the QSAR models created for the water treatment processes. DEFINITE 3.1 ver.3.1.1.7 software was used to perform the MCA [10]. The different steps involved in developing a DSS based on MCA are listed in *Figure 4.1*.

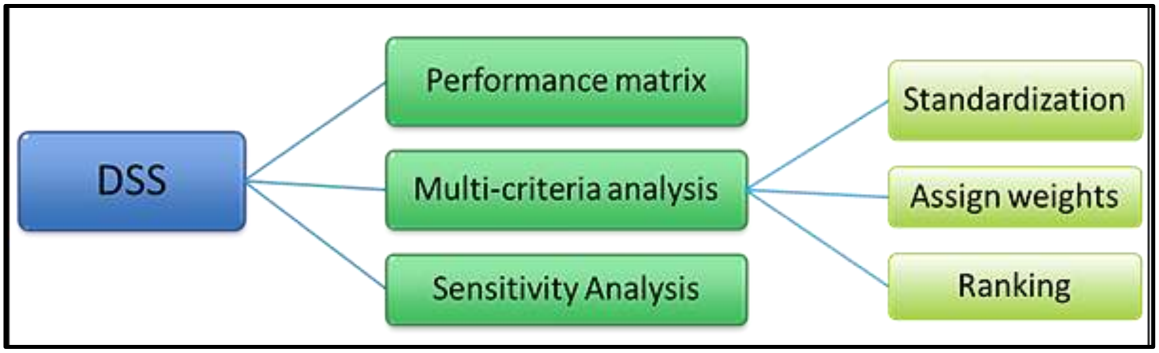


Figure 4.1: Steps involved in building a DSS

4.3. THEORY

4.3.1. Performance Matrix

The input data (i.e., scores) to an MCA include information on all alternatives (different water treatment processes) and all criteria including sub-criteria (socio-economic-technical aspects, *see section 4.4.2*). The alternatives are scored against the criteria in a performance matrix, which is shown in *Table 4.2*. The considered alternatives in this study distinguished between natural systems (RBF, CW) and advanced treatment processes which included the main membrane processes (RO, NF), adsorption by granular activated carbon (GAC), and oxidation by direct ozonation (O_3) as well as advanced oxidation processes (AOP) which apply ozone in combination with hydrogen peroxide (O_3-H_2O_2, or simply AOP) and ultraviolet (UV) light in combination with H_2O_2 (UV-H_2O_2 or UV-AOP).

4.3.2. Multi-Criteria Analysis (MCA)

4.3.2.1. Weighted Summation Method

The weighted summation method was chosen for this study since it is transparent, involves expert-opinions and is well-recognized for decision-making. The overall performance (preference) score for each alternative is the sum of the alternative's score for each criterion multiplied by the weight for that criterion. The formula used for weighted summation is:

$$\text{Score}\,(a_j) = \sum_{i=1}^{N} w_i * \hat{s}_{ij} \tag{1}$$

Where:
A is the set of alternatives with a_j (j = 1...M),
C is the set of effects with c_i (i=1…N),
s_{ij} is the score of alternative a_j for effect c_i,
$\hat{s}_{ij}$ is the standardized score of alternative a_j for effect c_i,
w_i is the weight of effect

In order to apply the weighted summation method, standardization and the assigning of weights are necessary. The final result is a ranking of the alternatives based on the overall score for each alternative.

4.3.2.2. Standardization

Scores with different measurement scales (percent-removal, $/m^3, kg/m^3) cannot be compared to each other directly. Therefore, the scores must be standardized to a dimensionless value between 0 and 1 before the overall score for each alternative can be calculated. In DEFINITE 3.1 software, several standardization functions are available. In this study, two linear standardization functions were used: goal standardization for quantitative criteria and maximum standardization for qualitative criteria (*Table 4.1*). In goal standardization, an ideal or goal value (best value) and a baseline value (worst value) are specified. Quantitative criteria (costs, carbon footprint) are well suited for this method since they usually have a baseline and an ideal value. The qualitative criteria were standardized by maximum standardization. Apart from the qualitative/quantitative type, the criteria are also classified into cost or benefit. For 'cost' criteria, which have a negative correlation between score and effect (e.g., for carbon footprint, the lower the better). 'Cost' criteria (negative impacts) thus have a minimization effect (a standardized line sloping downwards). For 'benefit' criteria which have a positive correlation between score and effect (e.g. percent removal, the higher the better). 'Benefit' criteria (positive impacts) thus have a maximization effect (a standardized line sloping upwards).

Table 4.1: Types of standardization methods: Maximum and Goal

Method	Characteristics	Formula
Maximum	scales the performance between minimum and maximum value	Benefit = $\frac{score}{highest\ score}$ Cost = $1 - \frac{score}{highest\ score}$
Goal	specify the highest and lowest value	Benefit = $\frac{score - baseline}{goal\ value - basline}$ Cost = 1 - $\frac{score - baseline}{goal\ value - baseline}$

4.3.2.3. Assigning Weights

In the MCA, weights are assigned to the different criteria to reflect different stakeholder perspectives and the final ranking is heavily influenced by these weights. A survey questionnaire was sent to two groups (academics (30 respondents) and industry professionals (8 respondents)). The weights were derived by the expected value method and entered into the DEFINITE software accordingly. The expected value method is based on ordinal weights wherein the decision maker ranks the criteria from most to least important or in certain cases gives equal ranks since they are equally important [10, 11]. In the survey, the main group of criteria (treatability, costs, technical, sustainability and time) and sub-groups were ranked from highest to lowest importance or equally ranked based on the personal opinions of participants. The average rank was calculated for each criterion within each main group and sub-groups, and the criteria were ranked accordingly. The rank orders of criteria were converted to quantitative weights (w) using the formula:

$$\mathrm{w_k} = \frac{1}{K}\sum_{i=k}^{K}\frac{1}{i} \qquad where\ K\ is\ the\ number\ of\ criteria \qquad (2)$$

The sum of the weights for the criteria/sub-criteria within each group is 1.0. Detailed information regarding the expert-based and neutral-perspective surveys is given below and in *section 4.4.3*:

i) *Expert-based survey*: A concise survey consisting of six questions (*see Appendix C*) was sent to researchers (30 PhD research students and post-doctoral research associates) in the Water Desalination and Reuse Center at King Abdullah University of Science and Technology (KAUST) providing an academic perspective. The survey was also sent to 8 professionals/operators from reputed international water-related companies (consultants, operators, and manufacturers) providing an industrial perspective. All the respondents of the survey were working in water-related fields and had good information on OMPs and the treatment processes associated to remove them. The responses from the surveys were compiled separately and the weights were assigned.

ii) *Neutral-perspective*: Here all of the criteria were assigned equal weights, i.e., no preference was given to any criterion over another and the MCA was performed.

4.3.2.4. Ranking

This step displays the results of the MCA in terms of ranks based on *equation 1*, i.e., the alternatives are being ranked based on the overall score.

4.3.3. Sensitivity Analysis

Sensitivity analysis helps to investigate which changes in the scores or weights are necessary to bring about a significant change, particularly if two alternatives have only a small difference in their overall score. A ranking is considered robust when it is not sensitive to variations in the scores or weights. In this study, the sensitivity analysis was performed for the scores of the quantitative criteria, i.e., unit costs and carbon footprint, and the assigned weights for the main groups: treatability, costs, technical, sustainability and time.

4.4. RESULTS AND DISCUSSIONS

4.4.1. Percent-removal

(a) MCA-I

In *Table 4.2,* three criteria (percent-removal, costs, carbon footprint) are quantitative and the remaining eleven criteria are qualitative terms. Since this study consisted of a wide range of alternatives and criteria, it was very difficult to find representative quantitative values for all the effects in the literature, hence qualitative terms were used where sufficient quantitative data were not available. The percent-removals are constant in MCA-I and an arbitrary value (99%) was chosen which does not influence the final ranking.

(b) MCA-II

In MCA-II, the percent-removals for a selected group of OMPs were experimental values taken from a published report in literature [1] . The percent-removals are shown in *Table 4.3*. The main differences between MCA-I and MCA-II is that the variable percent-removals in *Table 4.3* replace the fixed "percent-removals" in *Table 4.2,* and studies on CW are not present, all the other rows are the same as in *Table 4.2*.

(c) MCA-III

The percent-removals for the QSAR-based MCA were taken from QSAR models built for reverse osmosis [12], nanofiltration [13] , adsorption [14] , ozone [15, 16], advanced oxidation [16], and river bank filtration (RBF) as represented by soil aquifer treatment (SAT) [3]. In the case of QSAR models some compounds were either over- or under-predicted compared to the experimental values. The over/under predicted values were rounded off to ±20 of the observed values (e.g., for carbamazepine the observed value: 76% and QSAR predicted value: 53%; rounded off value was 76-20= 56%). The predicted values of more than 100% were rounded off to 100% (for e.g., naproxen the observed value: 99% and QSAR predicted value: 107%; rounded off value was 100%). The QSAR-predicted values are shown in *Table 4.4* and the modified (rounded off) values are highlighted in *bold*.

4.4.2. Other Criteria

(i) Natural Organic Matter (NOM) and By-product (BP) Formation

Some of the selected pretreatment alternatives considered in this study, i.e., oxidation, adsorption, and membrane processes, are greatly affected by NOM [17-19]. By-product formation is a major concern for the oxidation processes. In case of ozonation, bromate is the major concern since it is carcinogenic [20, 21]. As shown in *Table 4.2*, qualitative scores (---/0) are used to describe the effect of NOM on processes (e.g., “**---**“describes a large negative effect (NF, RO) and “0” no effect (RBF, CW).

(ii) Costs

The unit costs are in US Dollar per cubic meter of product water (\$/m^3). Cost is variable in time, site specific and expressed in different currencies of which exchange rates vary. Therefore, cost data produced by the same source should ideally be used for comparability reasons. In this study, efforts were made to obtain cost-data from the same source, however, due to limitation of data availability, not all cost-data is directly comparable. The sensitivity analysis was performed to analyze the fluctuations in the overall rankings due to the cost data. Costs were expressed as annualized value, combining the capital costs and the operational/maintenance costs (O/M costs). The size of the water treatment plants ranged from 33,000 - 99,000 m^3/day. The costs for river bank filtration (RBF) [22, 23], nanofiltration (NF) [24] and reverse osmosis (RO) [25] were taken from journal articles, while data for ozonation [26], advanced oxidation ([27], ultra-violet based advanced oxidation [27]), constructed wetlands [28] and adsorption [27] were taken from cost-related text books.

(iii) Technical Aspects

The technical aspects included in the criteria are reliability/maintenance, professional skill required and potential for modification. Reliability/Maintenance indicates the process stability and maintenance requirement. RBF is considered to be most reliable with least maintenance, followed by CW, ozonation, adsorption and membranes. Natural systems (RBF,CW) are easy to maintain compared to the advanced treatment systems such as membrane processes (RO,NF), which require regular backwashing, replacements etc. Systems which are more sophisticated such as membranes and ultra-violet-oxidation (UV-AOP) also require higher professional skill unlike natural systems, which are easier to operate [29]. In the case of PM, it is easier to retrofit or upgrade existing advanced processes than natural systems which can have hydro-geological issues [29].

Table 4.2: Performance matrix

Cost (C)/Benefit (B)		**Natural Systems**		**Adsorption and Membranes**			**Oxidation**			
TREATABILITY	**C/B**	**RBF**	**CW**	**GAC**	**NF**	**RO**	**O_3**	**AOP**	**UV-AOP**	**Ref.**
%removal*	**B**	99	99	99	99	99	99	99	99	
Natural Organic Matter		0	0	**--**	**---**	**---**	**--**	**--**	**--**	[17-19]
By-Products formation		0	0	0	0	0	**--**	**--**	**-**	[20, 21]
COSTS										
Unit Costs ($\$/m^3$)	**C**	0.06	0.02	0.132	0.234	0.53	0.114	0.114	0.137	[22-28]
TECHNICAL										
Reliability/ Maintenance		**++**	**+**	**+**	**--**	**--**	**-**	**-**	**-**	[29]
Professional Skill		**+**	**-**	**-**	**--**	**--**	**-**	**-**	**--**	[29]
Potential for Modification		**-**	**-**	**+**	**+**	**+**	**++**	**++**	**++**	[29]
SUSTAINABILITY										
Environ Impact										
Land		**--**	**--**	**+**	**+**	**+**	**++**	**++**	**++**	[29]
Waste discharge		0	0	**-**	**--**	**--**	0	0	0	[29]
Carbon footprint ($kgCO_2/m^3$)	**C**	0	0	0.42	0.41	1	0.053	0.053	0.053	[29, 30]
Resource Use										
Energy Use		**++**	**++**	**-**	**-**	**--**	**+**	**+**	**-**	[29]
Chemical Use		**++**	**+**	**-**	**--**	**--**	**-**	**--**	**-**	[29]
Public Acceptance		**++**	**-**	**++**	**++**	**++**	**++**	**++**	**+**	[2, 29]
TIME										
Residence time		**---**	**---**	**+**	**++**	**++**	**+++**	**+++**	**+++**	[29]

SCORES	***ALTERNATIVES***	***CRITERIA***
0: no effect	RBF: river bank filtration	%removal*: variable in MCA-II, MCA-III
---: large negative effect	CW: constructed wetlands	
--: intermediate effect	GAC: granular activated carbon	
-: small negative effect	NF: nano-filtration	
+++: small positive effect	RO: reverse osmosis	
++: intermediate effect	O_3: ozonation; UV/H_2O_2: ultra violet advanced oxidation	
+: small positive effect	O_3/H_2O_2: advanced oxidation	

Table 4.3: Pilot–scale percent removals of OMPs based on Snyder's report [1]

	RBF	GAC	NF	RO	O_3	AOP	UV-AOP
Acetaminophen	99	80	50	99	99	99	97
Androstenedione	99	85	80	99	99	99	96
Caffeine	98	80	80	99	97	97	89
Carbamazepine	13	80	80	99	99	99	88
DEET	91	70	80	95	76	82	89
Diazepam	65	70	80	95	82	85	93
Diclofenac	99	45	80	99	99	99	98
Dilantin	22	45	80	99	86	88	97
Erythromycin	98	80	85	99	92	92	64
Estradiol	99	85	80	99	99	99	98
Estrone	99	80	80	99	99	99	99
Ethinyl estradiol	99	95	80	99	99	99	99
Fluoxetine	99	85	85	99	99	99	99
Gemfibrozil	99	45	80	99	99	99	95
Hydrocodone	99	80	80	99	99	99	99
Ibuprofen	99	50	80	99	87	88	94
Iopromide	95	50	85	95	61	58	91
Meprobamate	74	50	80	85	59	60	75
Naproxen	98	50	50	99	99	99	99
Oxybenzone	97	90	85	99	99	99	66
Pentoxifylline	99	80	80	99	99	99	90
TCEP	32	40	80	99	8	9	16
Triclosan	98	90	85	99	99	98	97
Trimethoprim	99	80	80	99	99	99	94
Range	86	55	35	14	91	90	83
Mean	86.21	70.21	78.54	97.92	88.88	89.29	88.42

(iv) Sustainability

Sustainability of the water treatment processes was measured by the sub-group criteria, environmental impact (land use and waste discharge, carbon footprint), resource use (energy and chemical use) and public acceptance.

In case of land use, the amount of land space occupied for running the treatment process was considered. The natural systems (RBF, CW) scored worst in this regard due

Table 4.4: QSAR model values (% removals)

	RBF	**GAC**	**NF**	**RO**	**O_3**	**AOP**	**Reference**
Atrazine	no data	41	100	no data	**65**	67	[13, 14, 16]
Caffeine	**100**	no data	85	**45**	85	79	[3, 12, 16, 31]
Carbamazepine	18	**56**	93	78	94	**100**	[3, 13, 14, 16, 31]
Diclofenac	**100**	32	**100**	no data	**100**	**100**	[3, 13, 14, 16]
Estradiol	no data	92	96	99	94	90	[12-14, 16]
Estrone	no data	91	100	no data	97	94	[13, 14, 16]
Gemfibrozil	73	57	100	no data	90	93	[3, 13-16]
Ibuprofen	84	**42**	98	**62**	80	93	[3, 12-14, 16]
Lindane	no data	no data	82	98	7	15	[12, 13, 16]
Naproxen	**100**	53	**100**	no data	96	84	[3, 13, 14, 16]
Pentoxifylline	**100**	73	100	no data	91	88	[3, 13, 14, 16]
Sulfamethoxazole	no data	20	96	no data	87	89	[13, 14, 16]
Testosterone	no data	100	no data	95	85	96	[12, 14, 16]

to high land usage compared to advanced processes such as membrane or oxidation processes which are more compact. For waste discharge, the membrane processes were least preferred due to high concentrate, backwash and cleaning solutions, compared to oxidation and natural systems which do not have associated waste discharge [29].

The amount of carbon dioxide emitted from the processes is a major environmental concern, due to its high negative impact on climate change. The carbon footprint is usually defined as the total of the GHGs (greenhouse gases) produced from operational and embodied emissions. In this study, the carbon footprint was measured only in terms of CO_2 - emissions (kilograms CO_2/ m^3). The natural systems were preferred due to low CO_2 emissions, followed by oxidation, adsorption, and finally membranes which release considerable amount of CO_2 into the environment [29, 30].

Energy use refers to the power consumption required to run the processes. The advanced processes (RO, NF, UV-AOP, GAC) are far more energy intensive than the natural systems (RBF, CW). Membranes require energy for generating the desired pressure and backwashing, adsorption for regeneration, and oxidation for electrical generation of ozone. In the case of natural processes, the power consumption is least. Chemical use implies the use of chemical agents. These include anti-scalants in case of membranes, ozone and peroxide in case of oxidation, etc. In this case, the natural systems scored best due to least chemical usage and the membrane processes scored worst due to heavy use of chemicals [29].

Public acceptance reflects how well (or poor) a process is accepted by the public as a water treatment process to remove OMPs. Most of the processes are already well-accepted except CW which is slowly being considered for OMP removal [2, 29].

(v) Residence Time

The advanced processes, mainly oxidation, scored better since they remove the OMPs rapidly as compared to natural systems which require a longer residence time [29].

4.4.3. Weights

(i) Academic and Industrial Survey

The results of the survey are shown in *Figure 4.2*. Both the academic and industrial groups had the same opinion about the sustainability criteria, i.e., environmental impact (land use, waste discharge, and carbon footprint), resource use (energy use and chemical use), and public acceptance: Environmental impact was given highest priority, followed by resource use, while public acceptance was considered least important. Concerning the sub-criteria of environmental impact, waste discharge was considered the most

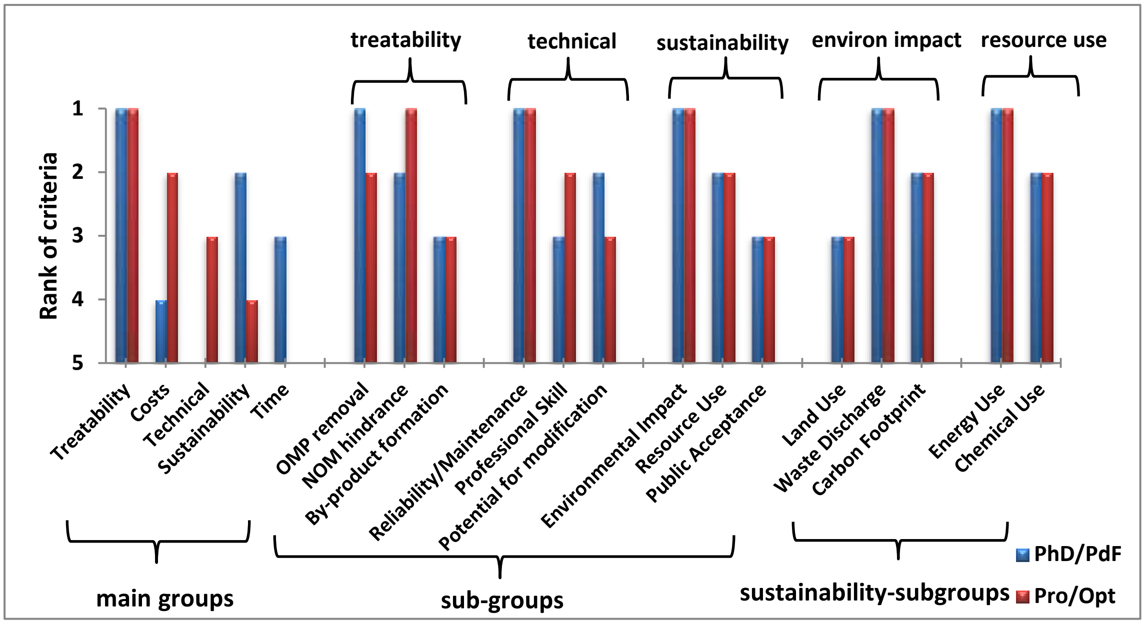


Figure 4.2: Academic vs Industrial survey rank comparison
Comparison of the preferences (ranks) given to the various criteria in the academic (researchers) and industry (professionals/operators (Pro/Opt) survey

important problem followed by carbon footprint and land use as the least important. In the case of resource use, both groups felt that energy use, in terms of power consumption, was more important than chemical use.

Concerning technical aspects, both groups considered reliability/maintenance to be the most important criterion. However, there was a difference in opinion for the criteria "professional skill" and "potential for modification": while academics considered potential for modification more important the industry representatives felt professional skill was more important.

Concerning treatability (i.e., percent-removal, NOM hindrance, and by-product formation), both academics and professionals considered by-product formation the least important. The academics gave highest priority to the criterion percent-removal whereas the professionals considered NOM hindrance to be most important. It is believed that professionals prioritized NOM hindrance because NOM presence adversely impacts most treatment processes by fouling, side-reactions etc. and therefore NOM removal is important for maintaining good operating conditions and hence for OMPs removal.

It was difficult to compare the two groups with regards to the main-criteria because the "time" criterion was included only in the survey among academics. However, from the five main criteria, the academic-group gave highest priority to treatability, followed by sustainability, time, costs and technical aspects. From the four main criteria, the professionals' decreasing order of preference was treatability followed by costs, technical aspects, and sustainability. The weights derived from the survey (according to equation 2) and which have been used in the MCA to reflect the preferences of the academic and professionals groups are listed in *Table 4.5.*

4.4.4. MCA-I: Ranking results

The objective of MCA-I was to compare and rank the water treatment processes based on all other criteria excluding percent-removal, therefore, the percent-removal rates of the OMPs were assumed to be identical (*Table 4.2*). *Figure 4.3* shows the MCA rankings (line chart, top) based on the weights (pie charts, bottom) that were derived from the academics and professional surveys and by assuming equal importance of all criteria (neutral perspective). In case of academics' and professionals' perspectives, RBF ranked highest among the different alternatives. The intermediate positions were occupied by oxidation (O_3, AOP, UV-AOP), adsorption (GAC) and CW. The membrane processes (RO, NF) always remained at the lower-end of the ranking since they are less environment friendly, more costly, and sophisticated to operate compared to the other processes. In case of neutral perspective, oxidation was the preferred process followed by RBF, adsorption (GAC), CW and membrane processes.

Table 4.5: Assigned weights for criteria: academic vs industry

	Main group		Sub-group		Sustainability-subgroup	
	Academic	Industry	Academic	Industry	Academic	Industry
Treatability	0.455 (45%)	0.520				
%removal			0.610	0.280		
NOM hindrance			0.280	0.610		
BPs			0.110	0.110		
Costs	0.089 (9%)	0.270				
Technical	0.040 (4%)	0.150				
R/M			0.610	0.610		
Skill			0.110	0.280		
PM			0.280	0.110		
Sustainability	0.257 (26%)	0.060				
Environmental Impact			0.610	0.610		
Land Use					0.110	0.110
Waste discharge					0.610	0.610
Carbon footprint					0.280	0.280
Resource Use			0.280	0.280		
Energy Use					0.750	0.750
Chemical Use					0.250	0.250
Public Acceptance			0.110	0.110		
Residence time	0.158 (16%)					
***Overall weight* (∑)**	1.0	1.0	1.0	1.0	1.0	1.0

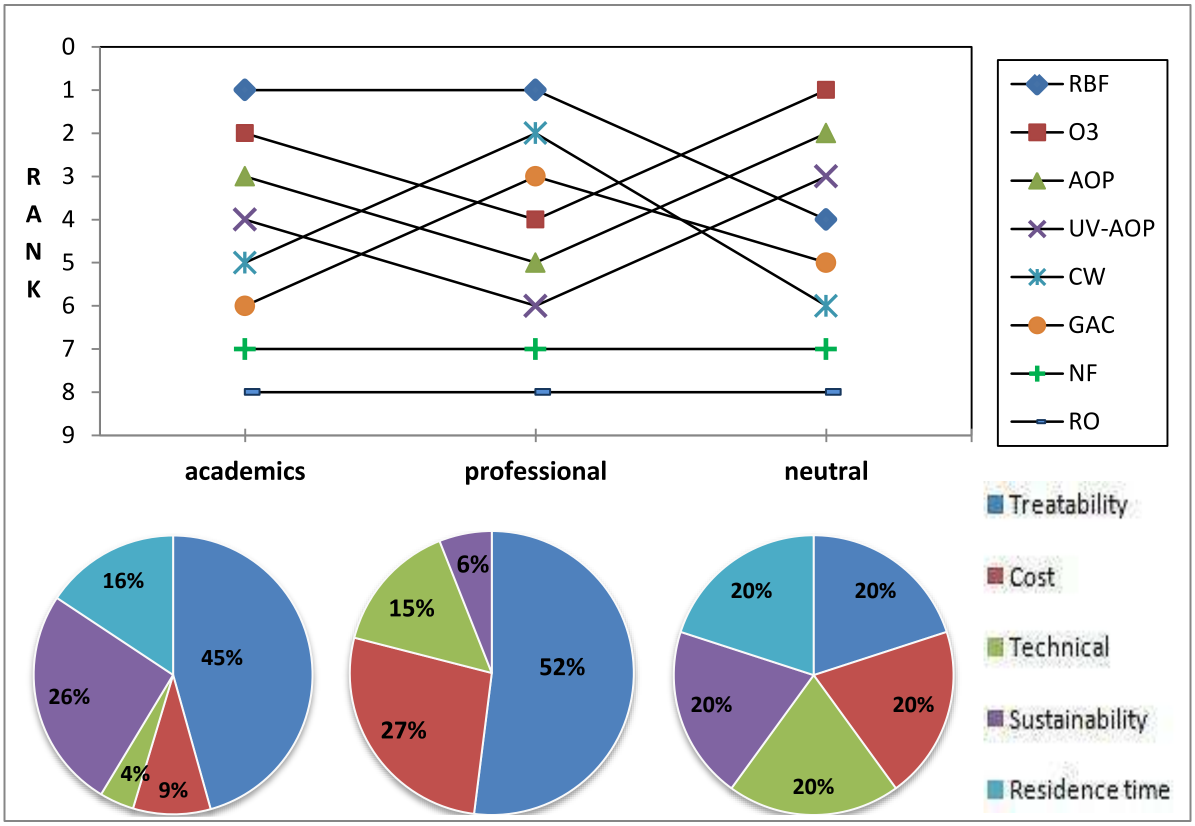


Figure 4.3: MCA-I process rankings
The rankings (line chart, top) based on scores in performance matrix (*Table 4.2*) and weights derived from survey (pie chart, bottom)

4.4.5. MCA-II: Experimental Study-Ranking

MCA-II is based on the experimental percent-removals of OMPs [1] as given in *Table 4.3* and the other criteria scores in *Table 4.2*. The weights for the criteria were taken from the academic survey. Twenty-four compounds were analyzed, for which RBF and oxidation (O_3, AOP, and UV-AOP) were found to be the preferred processes under this set of scores and weights in MCA-II, followed by adsorption and finally the membrane processes as shown in *Figure 4.4*. In case of DEET (N, N-Diethyl-meta-toluamide), an ozone-resistant compound, UV-AOP was found to be a better option compared to O_3 or AOP. For erythromycin, treatment by GAC was preferable over UV-AOP.

TCEP (tris (2-chloroethyl) phosphate) is the only compound in MCA-II for which membrane processes (RO and NF) outranked most other treatment alternatives, except for RBF which was again found to be the best alternative. However, RBF has a low TCEP

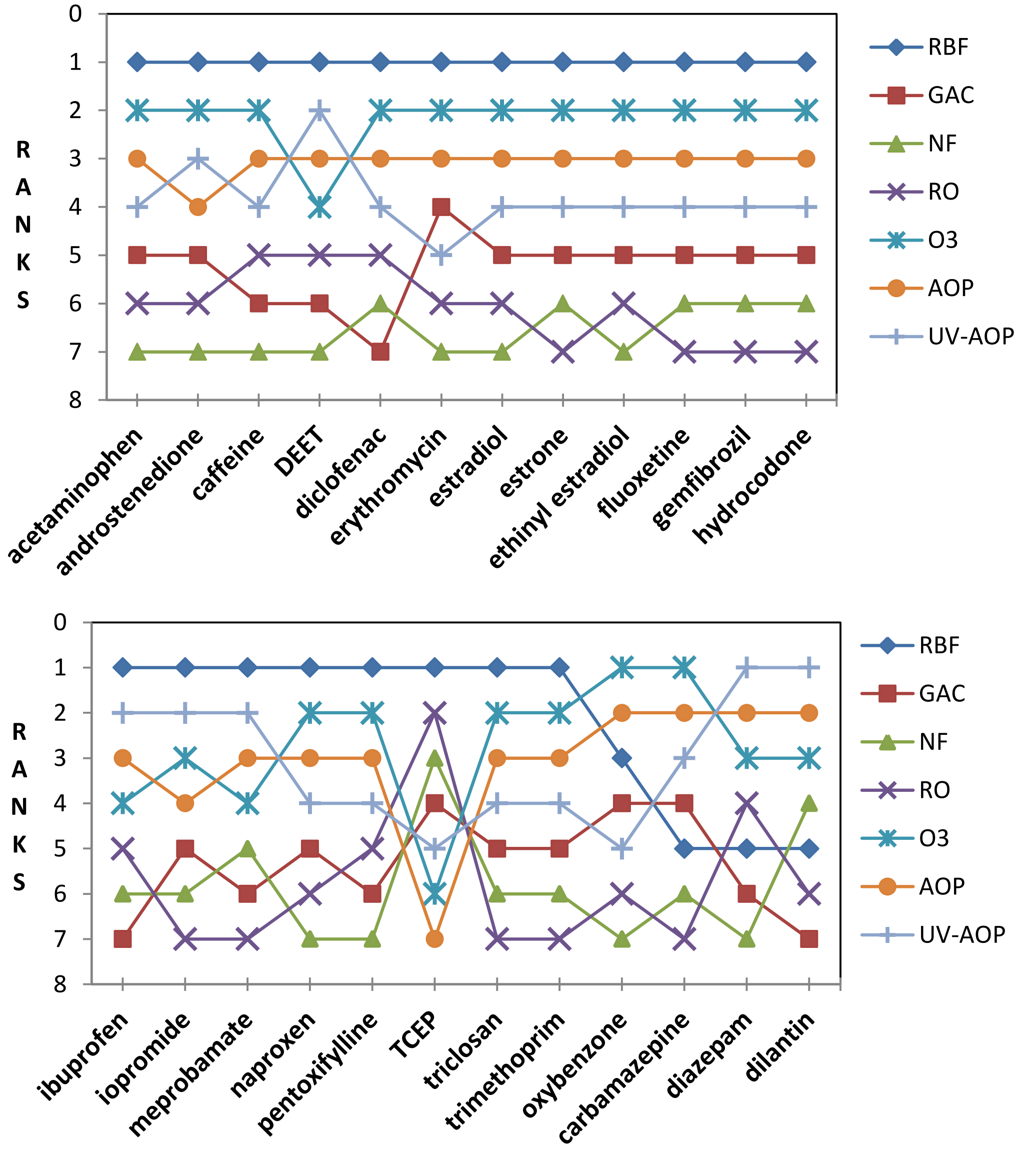


Figure 4.4: MCA-II process rankings
The rankings are based on experimental percent removals (*Table 4.3*) and the remaining criteria as shown in *Table 4.2*

removal rate of 32% and only performed best in the MCA because of a better performance in other criteria. If a minimum percent-removal rate, e.g. of 50%, is introduced as a cut-off criterion, RO (99% removal) and NF (80% removal) remain the only feasible options because all other alternatives have lower TCEP removal rates (*Table 4.3*).

In the case of oxybenzone, ozonation (O_3 and AOP) were better processes compared to RBF. Finally for carbamazepine, diazepam and dilantin, ozonation/oxidation (O_3, AOP and UV-AOP) were the best processes. In general, the decreasing trend was RBF > oxidation > GAC ≥ membrane, except for TCEP.

4.4.6. MCA-III: "QSAR-based Percent-Removal"-Rankings

The MCA-III rankings based on the QSAR predictive models (*Table 4.6*) built for OMPs removal were assessed for oxidation (direct ozonation [15, 16], advanced oxidation process [16]), membranes (nanofiltration [13], reverse-osmosis [12, 31]), adsorption [14] and RBF [3]. The QSAR-model equations are shown in *Table 4.6.* All of the QSAR equations had good performance indices and were mechanistically interpreted, validated.

In MCA-III, all the percent-removals for the OMPs are totally prediction based. The main advantage of MCA-III is that with further inclusion of QSAR models it can act as a good index to choose the best process for OMPs removal without any lab analysis, purely based on computational aspects. In *Figure 4.5*, the QSAR-based MCA-III is shown. For carbamazepine and ibuprofen, all six processes were compared. In general, for OMPs removal, oxidation (O_3, AOP) and RBF were the preferred processes. In case of lindane (chlorinated OMP), AOP was shown as the best process although removed by only 17%.

Table 4.6: QSAR equations

AOP

(ln) removal = $9.77 - 0.63\,(E_{LUMO}-E_{HOMO}) - 0.194\,(EA) + 0.02\,(\#\text{ring atoms})$

$r^2 = 0.902,\ q^2_{LOO} = 0.868$

Ozone

% removal = $67.3 + 0.0506\,(PISA) + 5.2\,(\#\text{metabolites}) + 4.34\,(\#rtvFG) - 0.114\,(WPSA)$

$r^2 = 0.858,\ q^2_{LOO} = 0.80$

(ln) removal = $12.45 - 0.95\,(E_{LUMO} - E_{HOMO}) - 0.32\,(MON)$

$r^2 = 0.893,\ q^2_{LOO} = 0.706$

RBF

%removal = $174.8\,(\#\text{imidazoles}) + 158.4\,(AR) - 98.1\,(\#CONN) - 1830.3\,(ME) + 1851.1$

$r^2 = 0.84,\ q^2_{LOO} = 0.64$

GAC

relative adsorbability = $0.2730\,({}^{8}X_p) + 0.00106\,(FOSA)$

$r^2 = 0.86,\ q^2_{LOO} = 0.82$

NF

% rejection = $265.15\,\text{eqwidth} - 117.36\,(\text{depth}) + 81.66\,(\text{length}) - 5.23\,(\log D) + 1348.09\,(SR) - 1447.82$

$r^2 = 0.75,\ q^2_{LOO} = 0.72$

RO

% rejection = $252.714\,\text{eqwidth} + 35.104\,\text{length} + 485.839\,SR - 590.714$

$r^2 = 0.882,\ q^2_{LOO} = 0.8477$

MON: mean oxidation number
#rtvFG: # of reactive functional groups
nCONN: # of urea (-thio) derivate
eqwidth: equivalent-width
ME: mean atomic sanderson electronegativity
PISA: Pi-surface area
EA: electron affinity
AR: aromatic ratio
SR: salt-rejection
${}^{8}X_p$: 8th-order simple-path chi index
WPSA: weakly polar component of the solvent accessible surface area
FOSA: hydrophobic component of the solvent accessible surface area
$E_{LUMO} - E_{HOMO}$: energy difference between lowest unoccupied and highest occupied molecular orbitals

Descriptors in RO-ANN based model
size of the smallest ring
molecular weight (Da)
dipole moment (Cm)
shape index kappa2
energy of the lowest unoccupied molecular orbital (eV)
dipole hybridization (Cm)

If a minimum percent-removal rate, e.g. of 50%, is introduced as a cut-off criterion, RO and NF remain the only feasible treatment options for lindane (*Table 4.4*). However, the rankings are influenced by the assigned weights for the criteria, which are based on the academic survey (in which the other four criteria, i.e., costs, technical aspects, sustainability and residence time represent 55% of the total weights whereas treatability accounts for 'only' 45% of the total). The ranking between AOP and NF varied only in the second decimal point (0.56, 0.55), however, when the weights for treatability are increased to 50%, NF outranks AOP. Hence in the case of lindane, NF can be considered as the best process rather than AOP.

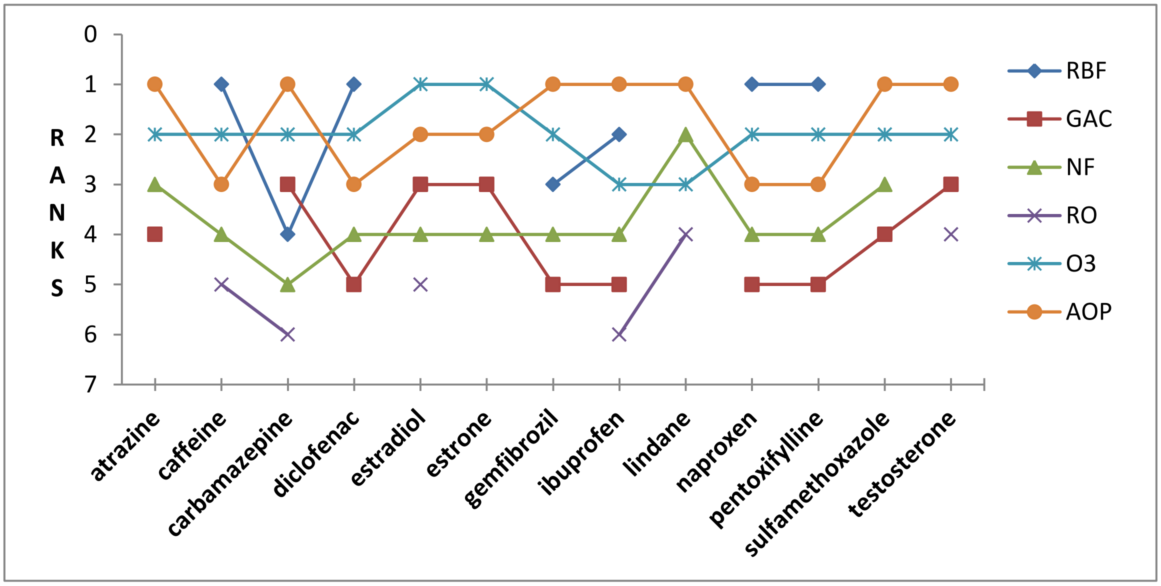


Figure 4.5: MCA – III process rankings.
The rankings are based on percent – removals from QSAR models (*Table 4.4*) and the remaining criteria as shown in *Table 4.2*

4.4.7. Ranking of three MCA studies by Neutral Perspectives

In this study, equal weights were assigned to all the criteria (20%) and only the percent removal of the OMPs were varied according to the study. MCA-I had fixed percent-removals of the OMPs, MCA-II had experimental values and MCA-III was based on the QSAR model values (*Figure 4.6*). In all the scenarios, natural systems and oxidation processes were found to be the best option, while membrane processes were found to be the least favorable option.

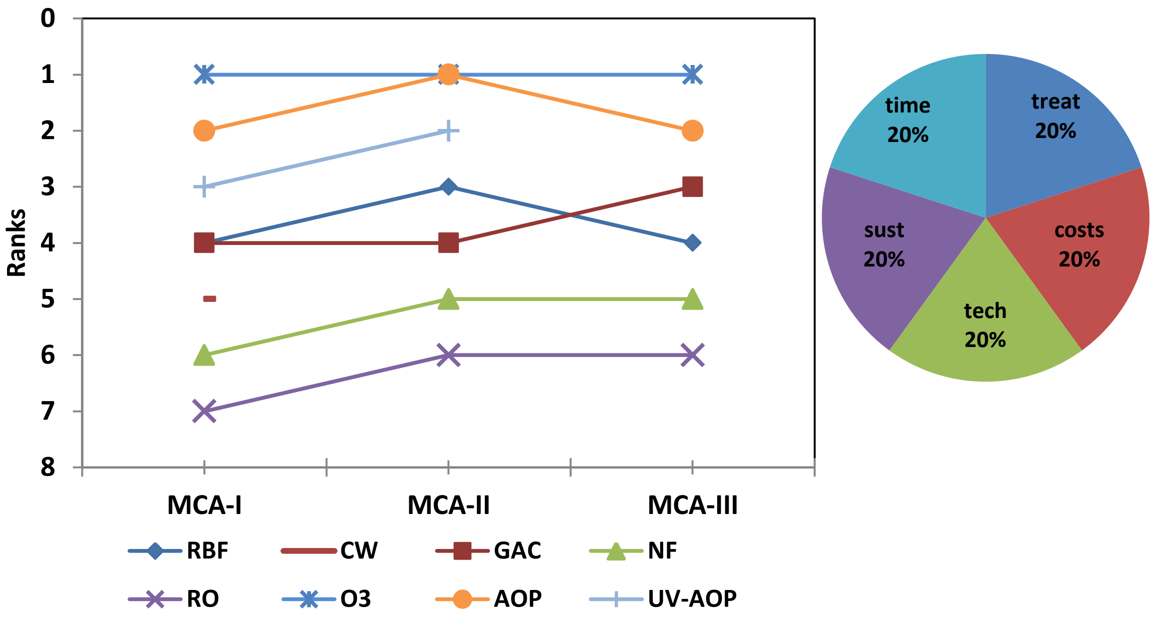


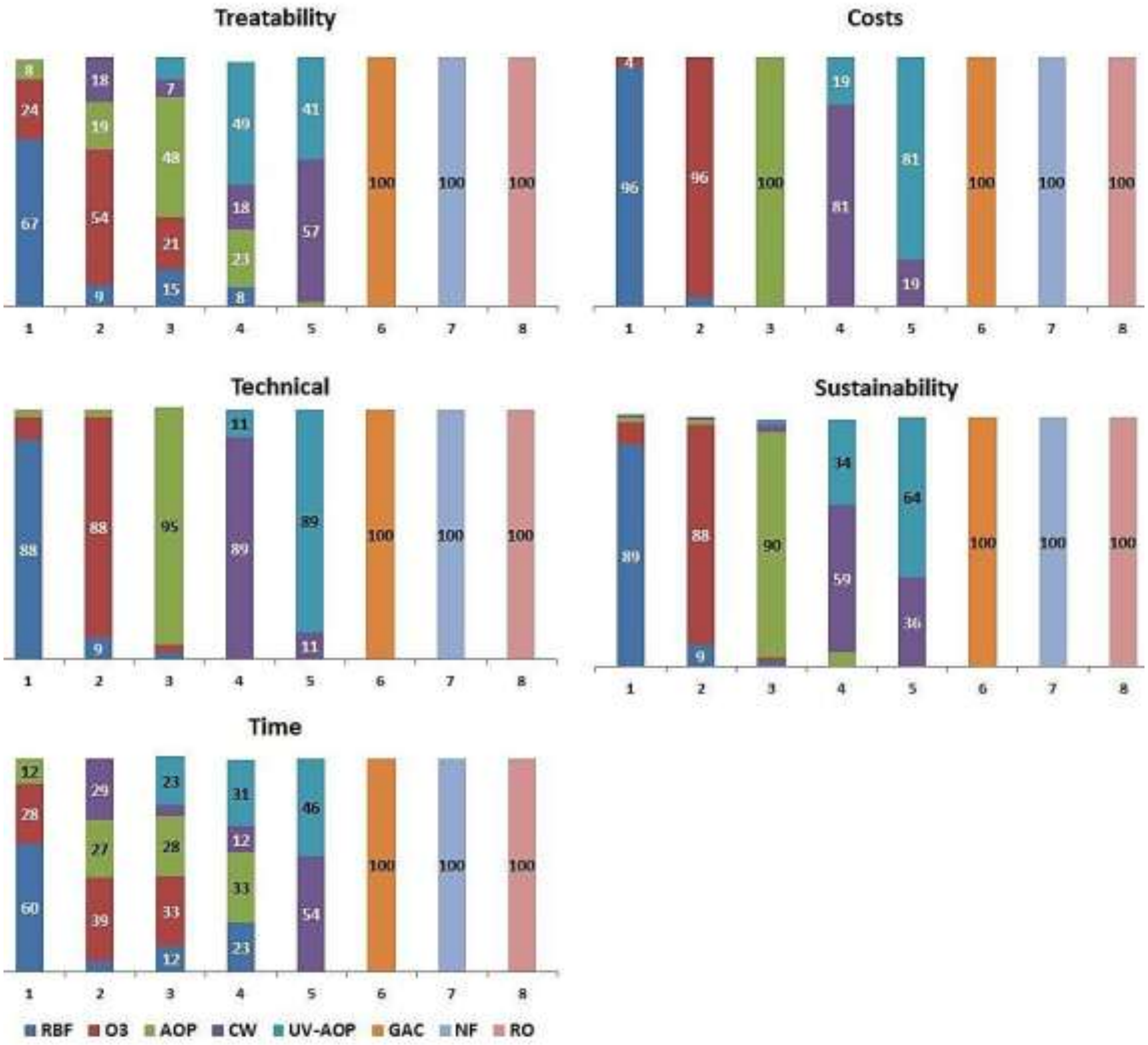
Figure 4.6: Neutral perspectives
The comparison of process ranks are based on experimental and fixed % removals

4.4.8. Sensitivity Analysis: Main Group Weights, Costs, and Carbon Footprint

A sensitivity analysis (50% - uncertainties) was performed for the weights of the main group criteria (treatability, costs, technical aspects, sustainability, and residence time) from the academic survey. There were no major changes in the rankings, i.e., RBF and oxidation processes were likely to rank highest whereas adsorption and membrane processes were likely to rank lowest, which is similar to the original MCA rankings (*Figure 4.7a*).

Second, the sensitivity of the ranking to changes in the scores of the cost criterion was investigated (*Figure 4.7b*). As observed, adsorption and membrane processes were not sensitive to changes in the cost-scores. In all cases, RBF had a high probability of about 60% to rank at the first position. Ozonation and constructed wetlands had probabilities of about 30% and 10% respectively, to rank at the first position, if the cost values were assumed to be up to 50% higher or lower than the original value. The oxidation processes (O_3, AOP, UV-AOP) and CW are likely to rank on second, third and fourth positions (probabilities between 55-65%) followed by adsorption and membrane processes, which had a high probability (100%) of ranking at positions six, seven and eight, respectively. This is generally similar to the original MCA rankings and indicates that the original MCA rankings are not very sensitive to variability in the cost data.

A sensitivity analysis was also conducted for carbon footprint, for which a similar trend as cost based sensitivity analysis was observed *(Figure 4.7c)*. There were no fluctuations in the rankings of adsorption and membrane processes. RBF ranked highest followed by ozonation, AOP, UV-AOP and CW even if 50% uncertainty in the carbon footprint data is to be assumed.



Treatability

Costs

Technical

Sustainability

Time

Figure 4.7a: Sensitivity analysis for main group criteria.

Sensitivity analysis assuming 50% uncertainties of the weights assigned to the main group criteria. The X-axis shows the rank of the alternatives (processes, 1-8); the Y-axis shows the percent probability that a treatment process occupies a certain rank. For example, if the treatability of RBF was assumed to be up to 50% higher or lower than the actual value assumed in this study, RBF has a probability of 67% to rank at first position, O_3 has a probability of 54% to rank second, etc. whereas GAC, NF and RO have a probability of 100% to rank at positions 6, 7 and 8, respectively

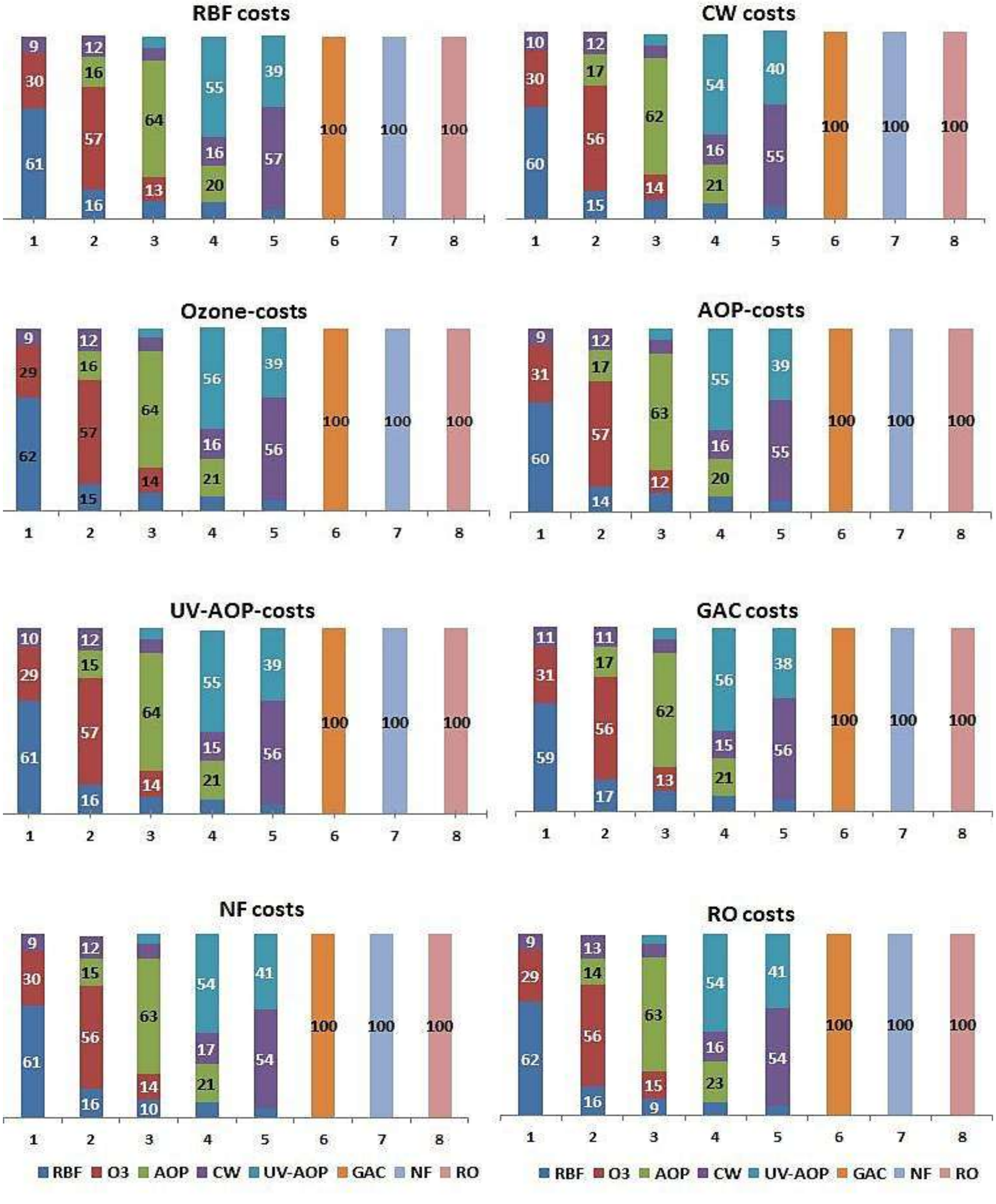


Figure 4.7b: Sensitivity analysis for cost data scores.

Sensitivity analysis assuming 50% uncertainty of the cost data.The X-axis shows the ranks of the processes (1-8); the Y-axis shows the percent probability that a treatment process occupies a certain rank. For example, if the cost value of RBF was assumed to be up to 50% higher or lower than the actual value assumed in this study, RBF has a probability of 61% to rank at first position, O_3 has a probability of 57% to rank second, etc. whereas GAC, NF and RO have a probability of 100% to rank at positions 6, 7 and 8, respectively

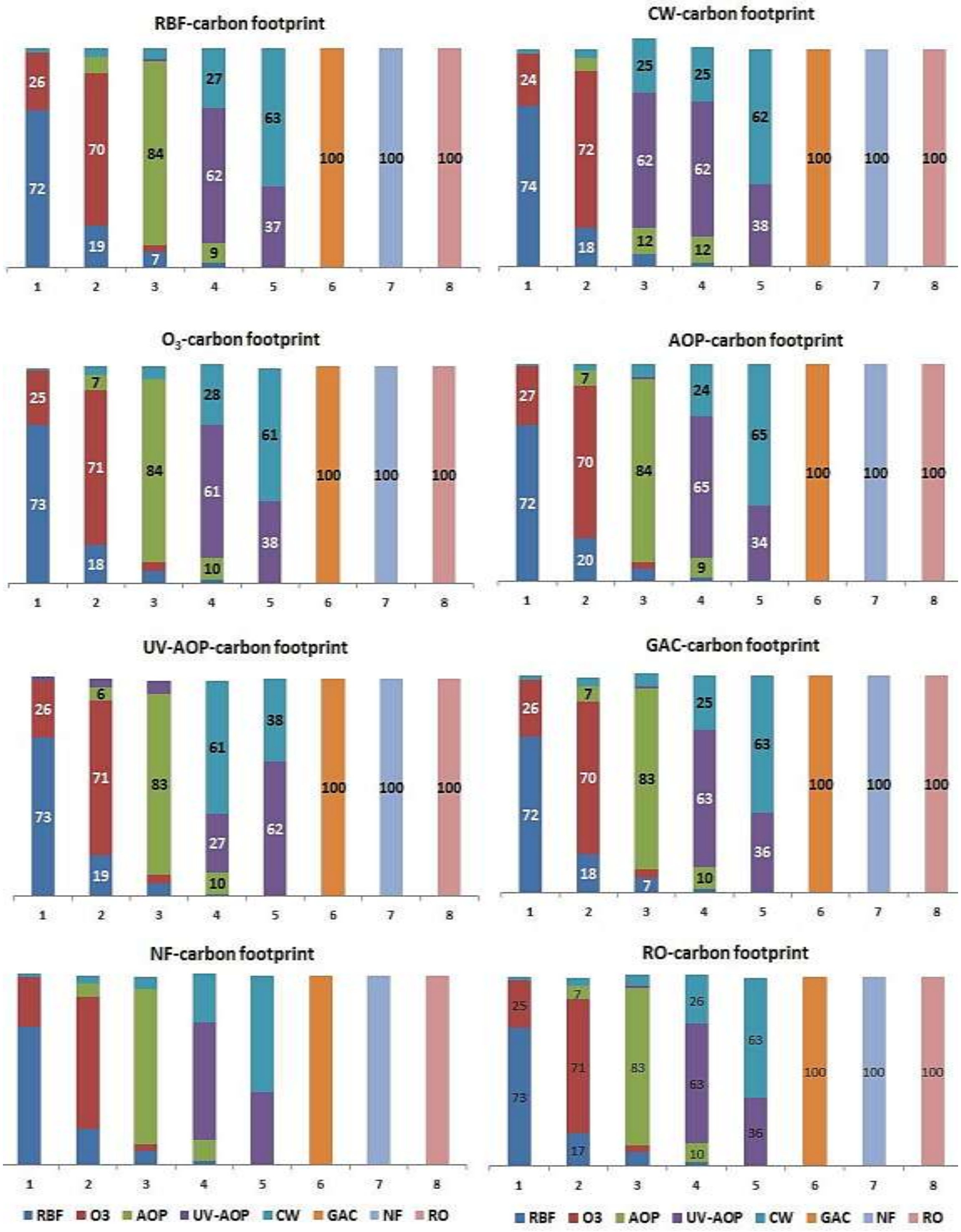


Figure 4.7c: Sensitivity analysis for carbon footprint scores

4.5. CONCLUSIONS AND RECOMMENDATIONS

In this study, a DSS with three types of MCA: MCA-I, MCA-II, MCA-III was created. MCA-I shows the ranking of the water treatment processes based on a set of technical and environmental criteria. MCA-II and MCA-III yield the process rankings with different experimental and QSAR-based percent-removals, respectively, based on the same technical and environmental criteria used in MCA-I.

In MCA-I, under the assumption of constant percent-removal rates for the OMPs, environment-friendly processes such as RBF were found to be preferable over advanced processes such as membranes or adsorption. In MCA-II, including the experimental percent-removal rates, the process rankings were more variable, reflecting the performance with regard to OMPs removal. Generally, RBF and oxidation were preferred alternatives compared to adsorption and membranes for the majority of the compounds. In case of TCEP, recalcitrant to oxidation, membranes were found to be the best option under the given set of criteria scores and weights. MCA-III was the most interesting and innovative, since it involved predicted values from validated QSAR models. Similar to the results of MCA-I and MCA-II, RBF and oxidation were the preferred treatment alternatives for most of the compounds, however, if the weight of the ‘treatability’-criterion is increased to more than 50% in the MCA; NF may become the best treatment option for some of the OMPs such as lindane. With the inclusion of more QSAR models for different treatment processes the DSS can be expanded with minimal experimental work.

The DSS would be helpful in the experimental set-up and planning of a drinking water treatment train for OMPs removal. The rankings obtained in this study are valid only for the given set of alternatives, criteria, scores and weights, and can be revised/refined, if new and better data are available.

Sensitivity analysis was performed for the quantitative scores (costs, carbon footprint) and the assigned weights for the criteria. The sensitivity analysis showed that the original MCA rankings are not very sensitive to a 50%-change in the scores/weights, i.e., RBF still has a high probability to rank first followed by oxidation, adsorption and membrane processes even if the cost and carbon footprint data are up to 50% higher or lower than the original values.

The DSS in this study is only for prioritization of the processes for OMP removal and not for decision making in specific projects for process design. Therefore, it is necessary to combine this DSS with a general decision making process for a system design, wherein more detailed design criteria would be evaluated. The DSS can be appropriately used for OMP removal by multi-barrier systems since water treatment systems consist of several consecutive unit processes. The DSS opens the window for a hybrid of natural systems (RBF) with advanced treatment processes (oxidation) for efficient OMPs removal, thereby reducing energy use and creating a greener, sustainable approach. For a site specific analysis, more detailed information about the hydro-geological conditions and site-based costs need to be considered.

REFERENCES

1. Snyder, S.A., Wert, E.C., Lei, H., Westerhoff, P., Yoon, Y., *Removal of EDCs and Pharmaceuticals in Drinking and Reuse Treatment Processes. Project # [2758]*. 2007, Awwa Research Foundation Denver, Colorado. p. 331.

2. Dordio, A., Carvalho, A.J.P., Teixeira, D.M., Dias, C.B., Pinto, A.P., *Removal of pharmaceuticals in microcosm constructed wetlands using Typha spp. and LECA.* Bioresource Technology, 2010. **101**(3): p. 886-892.

3. Maeng, S.K., *Multi-Objective treatment aspects of bank filtration*, in *Environmental Engineering and Water Technology*. 2010, TUDelft, UNESCO-IHE: Delft. p. 210.

4. Kiker, G.A.B., Todd S; Varghese, Arun; Seager, Thomas P;Linkov, Igor, *Application of multicriteria decision analysis in environmental decision making.* Integrated Environmental Assessment and Management, 2005. **1**(2): p. 95-108.

5. Stewart, T.J., Joubert, A., Janssen, R., *MCDA framework for fishing rights allocation in South Africa.* Group Decision and Negotiation, 2010. **19**(3): p. 247-265.

6. Belton, V.,Stewart, T., *Problem Structuring and Multiple Criteria Decision Analysis.* Trends in Multiple Criteria Decision Analysis, 2010: p. 209-239.

7. Janssen, R., *On the use of multi-criteria analysis in environmental impact assessment in The Netherlands.* Journal of Multi-Criteria Decision Analysis, 2001. **10**(2): p. 101-109.

8. Lattemann, S., *Development of an environmental impact assessment and decision support system for seawater desalination plants.* PhD Thesis, 2010: p. 183-224.

9. Lienert, J., Koller, M., Konrad, J., McArdell, C.S., Schuwirth, N., *Multiple-criteria decision analysis reveals high stakeholder preference to remove pharmaceuticals from hospital wastewater.* Environmental Science and Technology, 2011. **45**(9): p. 3848-3857.

10. Janssen, R.,van Herwinjnen, M., *DEFINITE [: a system to support decisions on a finite set of alternatives*. 1992: Kluwer Academic Publishers.

11. Rietveld, P., *Multiple objective decision methods and regional planning*. 1980: North-Holland Publishing Company.

12. Libotean, D., Giralt, J., Rallo, R., Cohen, Y., Giralt, F., Ridgway, H.F., Rodriguez, G., Phipps, D., *Organic compounds passage through RO membranes.* Journal of Membrane Science, 2008. **313**(1): p. 23-43.

13. Yangali-Quintanilla, V., Sadmani, A., McConville, M., Kennedy, M., Amy, G., *A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes.* Water Research, 2010. **44**(2): p. 373-384.

14. Redding, A.M., Cannon, F.S., Snyder, S.A., Vanderford, B.J., *A QSAR-like analysis of the adsorption of endocrine disrupting compounds, pharmaceuticals, and personal care products on modified activated carbons.* Water Research, 2009. **43**(15): p. 3849-3861.

15. Lei, H.,Snyder, S.A., *3D QSPR models for the removal of trace organic contaminants by ozone and free chlorine.* Water Research, 2007. **41**(18): p. 4051-4060.

16. Sudhakaran, S., Calvin, J., Amy, G.L., *QSAR models for the removal of organic micropollutants in four different river water matrices.* Chemosphere, 2012. **87**(2): p. 144-150.

17. Hong, S.,Elimelech, M., *Chemical and physical aspects of natural organic matter (NOM) fouling of nanofiltration membranes.* Journal of Membrane Science, 1997. **132**(2): p. 159-181.

18. Zhang, M., Li, C., Benjamin, M.M., Chang, Y., *Fouling and natural organic matter removal in adsorbent/membrane systems for drinking water treatment.* Environmental Science and Technology, 2003. **37**(8): p. 1663-1669.

19. Zularisam, A., Ismail, A., Salim, R., *Behaviours of natural organic matter in membrane filtration for surface water treatment--a review.* Desalination, 2006. **194**(1-3): p. 211-231.

20. Bonacquisti, T.P., *A drinking water utility's perspective on bromide, bromate, and ozonation.* Toxicology, 2006. **221**(2-3): p. 145-148.

21. Siddiqui, M.S., Amy, G.L., Murphy, B.D., *Ozone enhanced removal of natural organic matter from drinking water sources.* Water Research, 1997. **31**(12): p. 3098-3106.

22. Ray, C.G., T;Hubbs, S;Drewes, J;Haas, D;Darnault, C., *Riverbank Filtration for Drinking Water Supply.* 2008.

23. Sharma, S., Amy, G., Shaw, R. *Bank filtration: a sustainable water treatment technology for developing countries*. 2009: Water, Engineering and Development Centre (WEDC) Loughborough University of Technology.

24. Costa, A.R.,De Pinho, M.N., *Performance and cost estimation of nanofiltration for surface water treatment in drinking water production.* Desalination, 2006. **196**(1-3): p. 55-65.

25. Karagiannis, I.C.,Soldatos, P.G., *Water desalination cost literature: review and assessment.* Desalination, 2008. **223**(1-3): p. 448-456.

26. Chen, W., H; Haunschild, K.; Lund, J.R., *Delta Drinking Water Quality and Treatment Costs Technical Appendix H.* 2008.

27. Melin, G., *Treatment Technologies for Removal of Methyl Tertiary Butyl Ether (MTBE) from Drinking Water: Air Stripping, Advanced Oxidation Processes, Granular Activated Carbon, Synthetic Resin Sorbents: Prepared for the California MTBE Research Partnership by the National Water Research Institute.* Fountain Valley, CA, 2000.

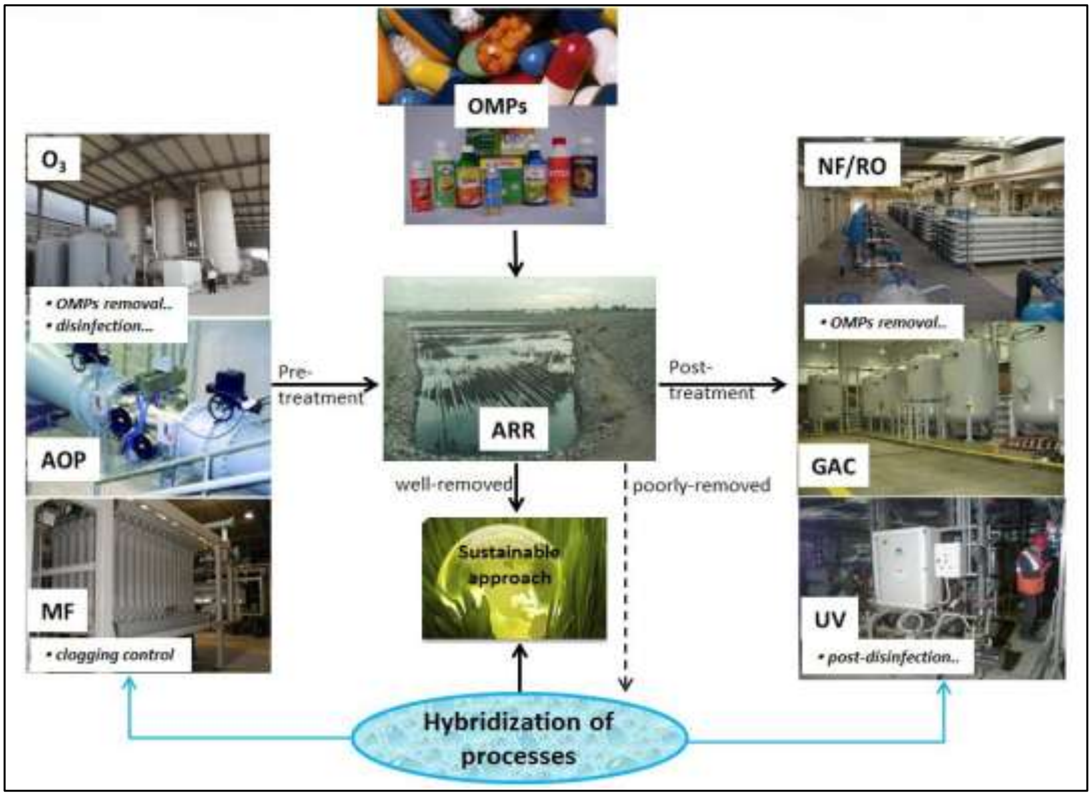
28. Kadlec, R.H.,Wallace, S., *Treatment wetlands*. 2009: CRC.

29. Crittenden, J.C., Trussell, R.R., Hand, D.W., Howe, K.J., Tchobanoglous, G., *MWH's Water Treatment: Principles and Design*. 2012: Wiley.

30. Oricawatercare, *Carbon footprint of TOC removal technologies: Comparison of the MIEX process, GAC, Ozone and High Pressure Membrane.* http://www.miexresin.com/index.asp, 2009.

31. Yangali-Quintanilla, V., Verliefde, A., Kim, T.U., Sadmani, A., Kennedy, M., Amy, G., *Artificial neural network models based on QSAR for predicting rejection of neutral organic compounds by polyamide nanofiltration and reverse osmosis membranes.* Journal of Membrane Science, 2009. **342**(1-2): p. 251-262.

5. HYBRIDIZATION OF NATURAL SYSTEMS WITH ADVANCED TREATMENTS FOR ORGANIC MICROPOLLUTANT REMOVALS: NEW CONCEPTS IN MULTI-BARRIER TREATMENT



HIGHLIGHTS

- Hybrids of natural and advanced treatment processes for OMPs removal
- Oxidation and microfiltration as pre-treatment to artificial recharge and recovery (ARR)
- Ultra, nanofiltration, adsorption, chlorination as post-treatment to ARR

This chapter is based on the following published paper and conference proceedings:

Sairam Sudhakaran, Sung Kyu Maeng, Gary Amy. Hybridization of natural systems with advanced treatments for organic micropollutant removals: New concepts in Multi-Barrier Treatment. *Chemosphere* 92 (2013) 731-737

Gary Amy, Saroj Sharma, Min Yoon, Sung Kyu Maeng, **Sairam Sudhakaran**. SIWW-2011-Singapore: Hybridization of Aquifier Recharge and Recovery (ARR): New concepts in Multi-Barrier Treatment for wastewater reuse

5.1. INTRODUCTION

As suggested in the previous chapter on DSS, a hybrid of a natural system (e.g., RBF or CWs) with an advanced treatment process (e.g., oxidation, membranes, or adsorption) can be implemented for efficient OMPs removal, thereby reducing energy use, carbon footprint and creating a greener, sustainable approach. In this chapter studies were focused on the hybridization of these processes.

Advanced water treatment processes such as reverse osmosis (RO), nanofiltration (NF), granular activated carbon (GAC) adsorption and oxidation, which includes direct ozonation (O_3) and advanced/ultra-violet oxidation process which involve $^{\bullet}OH$ radicals (AOP, UV-AOP), are industry choices for OMP removal; however, even natural systems such as riverbank filtration (RBF), aquifer recharge and recovery (ARR) as well as constructed wetlands (CWs) show good removals of most OMPs [1, 2]. Although the advanced treatment processes remove the OMPs efficiently, they are less sustainable because of high greenhouse gas emissions, high power consumption, and by-product/metabolite formation. Hybridization of advanced treatment processes with natural systems (ARR, CWs) can provide synergy in efficiently removing OMPs in a more sustainable manner. Several EU projects have focused on the removal of OMPs by natural and hybrid systems such as sustainable water management in the city of the future and water reclamation technologies for safe artificial ground recharge [3, 4].

In this study, hybrids of ARR coupled with an oxidation (O_3, AOP, and UV-AOP), membrane (NF/RO), or adsorption (GAC) process for OMP removal were studied. These hybrids provide a multi-barrier approach to eliminate bio-degradable and non-biodegradable compounds. *Figure 5.1* shows that OMPs which are well-removed by ARR obviously lead to a very sustainable approach. For compounds which are poorly biodegraded, the hybridization of processes can be a suitable alternative. Oxidation and microfiltration (MF) can act as pre-treatments to biodegradation. Oxidation helps in OMP removal and, in the case of ozone, also disinfection; MF helps to reduce ARR infiltration/recharge basin clogging [5]. Ultra-filtration (UF), NF, GAC and UV/chlorination can act as post-treatment to biodegradation. NF and GAC help to remove refractory OMPs, and UV/chlorination can serve as final disinfection. ARR as a post-treatment to oxidation helps to biodegrade the oxidation by-products/metabolites and serves as a pre-treatment to UF, NF, and GAC to minimize membrane fouling and dissolved organic carbon (DOC) loading.

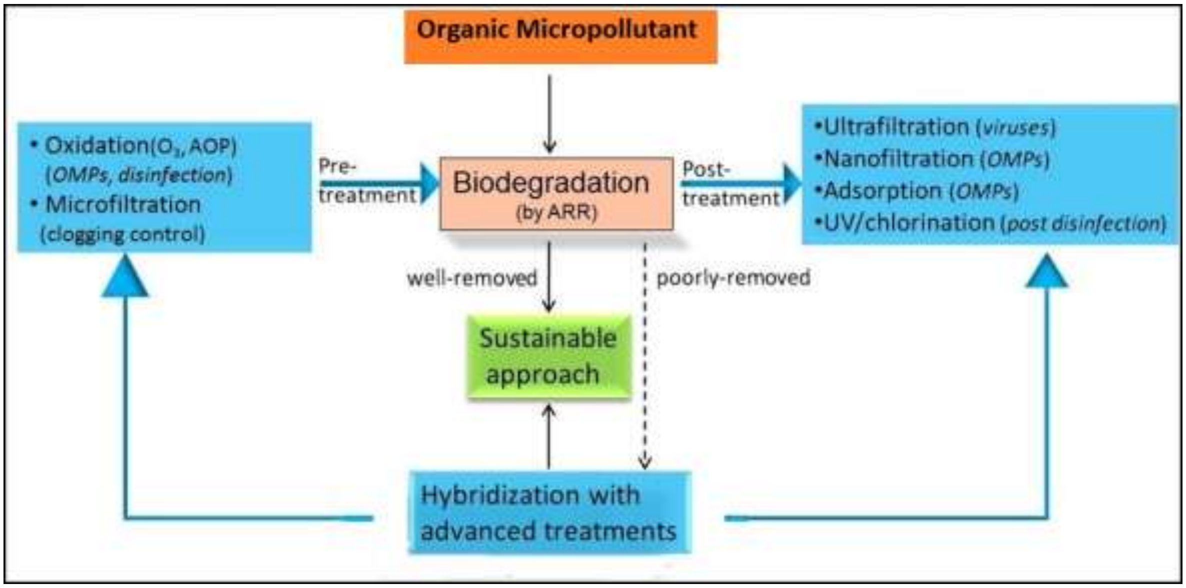


Figure5.1: Potential hybridization of ARR with other processes for OMP removal

5.2. THEORY

5.2.1 Natural treatment systems

All of the natural treatment processes such as RBF, ARR, or CWs have the same mode of mechanism for OMP removal, based on biodegradation. They are natural treatment processes which induce surface water to flow in response to a hydraulic gradient through soil sediments into a vertical or horizontal well. They are relatively low-cost, efficient and sustainable technology if the system is designed and operated with proper guidelines in an appropriate hydro-geological location. The factors that affect the performance of natural systems include: (i) raw water quality, (ii) well types, (iii) location and alignment of wells, (iv) residence time, (v) clogging, (vi) Schmutzdecke (bioactive layer), and (vii) redox conditions. These natural processes have shown potential to remove both natural organic matter and effluent organic matter that serve as a precursor to disinfection by-products [6, 7]; and pathogenic microorganisms such as *Cryptosporidium*, *Giardia* and viruses [8, 9] as well as OMPs [2].

5.2.2 Oxidation (O_3, AOP, UV-AOP)

O_3 is a good option to remove OMPs, since O_3 exhibits selectivity towards certain OMPs and readily transforms them. It reacts with double bonds, activated aromatic rings and specific ring atoms. In general, electron-pumping groups on a benzene ring enhance ozonation, and electron-withdrawing groups (nitro, halogens) decrease ozonation [10] . However, apart from the nature of the substituents, other parameters such as steric factors and connectivity between the aromatic rings need to be considered. O_3 is not very stable in water, it partly decomposes into $^{\bullet}OH$ radicals, leading to AOP which is less selective and more reactive than ozone, thereby oxidizing a wider range of OMPs with higher oxidation efficiency.

5.2.3 Adsorption (GAC)

GAC is also well-established for OMP removal. There are several properties that influence adsorption, namely hydrophobicity, charge, specific functional groups, molecule dissociation, polarity and polarizability [11]. Hydrophobic OMPs have low affinity towards water; when these OMPs are dissolved in water, they rearrange themselves in an energetically less favorable configuration and finally are adsorbed onto the adsorbent [12-14]. Charge effects can result in either attraction or repulsion of OMPs to an adsorbent surface. OMPs of an acidic nature release an H^+ ion and obtain a negative charge whereas OMPs of a basic nature obtain an H^+ ion and consequently a positive charge. The amount of H^+ dissociation/uptake depends on solution pH and the pK_a or pK_b of the OMPs. As a result, solution pH is a crucial factor in this mechanism, as both adsorbent and OMP charge depends on this parameter [11]. Electron-withdrawing functional groups on OMPs such as carbonyl and carboxyl groups reduce the adsorption ability as compared to electron pumping groups such as amines and aromatic rings [15, 16]. A major limitation of GAC is reduced loading caused by background NOM.

5.2.4 Membrane Separation (NF, RO, MF, and UF)

Membrane technologies are being increasingly used for OMP removal. They are the industry standard to remove a range of contaminants, however, membranes consume significant power and are less sustainable. Pressure-driven membrane processes, often used in water treatment, use hydraulic pressure to force water molecules through the membrane. Impurities are retained and concentrated in the reject, which becomes the reject water or concentrate stream. The permeate that passes through the membrane is recovered as product or pure water. The pressure-driven membranes, in order of decreasing permeability, are as follows: MF > UF > NF > RO. MF and UF are used to remove large organic molecules, colloids and micro-organisms. NF or RO, due to their small pore size, are used to remove OMPs and inorganic micropollutants such as arsenic and fluoride.

5.3. METHODOLOGY

Experimentally determined percent-removal of OMPs were taken from a research report [2]. The chemical structures of the OMPs are available in *Appendix A1*. SPSS version 17 was used to build quadrant plots. These quadrant plots are scatter plots with four quadrants. *Figure 5.2* illustrates the four quadrants. The horizontal and vertical axes represent the percent removal of OMPs by two different processes, usually ARR in hybrid with an advanced process (membrane, adsorption or oxidation). Quadrant I indicates poor removal of OMPs by both processes. Quadrant II and IV indicate good removal by one process and poor removal by the other. Quadrant III indicates good-removal by both processes. The experimental conditions for the water treatment processes are summarized in *Table 5.1*. The percent-removal of OMPs used in the quadrant plots were pilot-scale experimental results (*Table 5.2*).

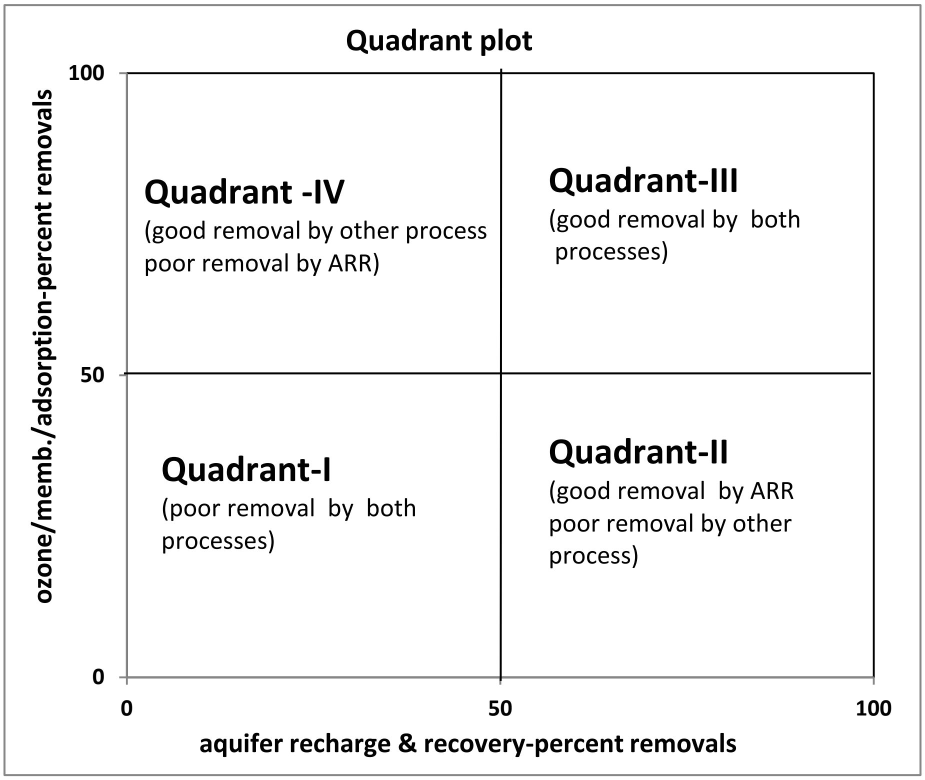


Figure 5.2: Quadrant plots with description of the quadrants

Table 5.1: Different conditions associated with different water treatment processes

	Residence time	Dosage	Membranes	GAC
ARR	36 days	not applicable	not applicable	not applicable
Ozone (O_3)	24 min	2.5 mgL^{-1}	not applicable	not applicable
AOP (O_3/H_2O_2)	24 min	2.5:0.065 mgL^{-1}	not applicable	not applicable
GAC	Empty bed contact time: 7.6 min	not applicable	not applicable	Norit Americas, Hydrodarco 4000
NF	not applicable	not applicable	ESNA, Hydranautics	not applicable
RO	not applicable	not applicable	Koch, Saehan, Osmonics	not applicable

Table 5.2: Pilot-scale percent removals of OMPs for different treatment processes [2]

Compound	Abbreviation	ARR	GAC	NF	RO	O_3	AOP	UV-AOP
acetaminophen	ACT	99	85	50	99	99	99	97
androstenedione	AND	99	95	80	99	99	99	96
caffeine	CAF	98	85	80	99	97	97	89
carbamazepine	CARB	13	85	80	99	99	99	88
DEET (N,N-diethyl-3-methylbenzamide)	DEET	91	85	80	95	76	82	89
diazepam	DIAZ	65	80	80	95	82	85	93
diclofenac	DICLO	99	50	80	99	99	99	98
dilantin	DIL	22	50	80	99	86	88	97
erythromycin	ERY	98	85	85	99	92	92	64
estradiol	ESTR2	99	95	80	99	99	99	98
estriol	EST	99	85	80	99	99	99	99
estrone	ESTRO	99	95	80	99	99	99	99
ethinyl estradiol	ET-ESTR2	99	95	80	99	99	99	99
fluoxetine	FLX	99	85	85	99	99	99	99
gemfibrozil	GEM	99	50	80	99	99	99	95
hydrocodone	HYDRO	99	85	80	99	99	99	99
ibuprofen	IBU	99	50	80	99	87	88	94
iopromide	IOPRO	95	50	85	95	61	58	91
meprobamate	MEP	74	50	80	85	59	60	75
naproxen	NAPRO	98	50	50	99	99	99	99
oxybenzone	OXYB	97	95	85	99	99	99	66
pentoxifylline	PENT	99	85	80	99	99	99	90
TCEP(tris-2-chloroethylphosphate)	TCEP	32	40	80	99	8	9	16
triclosan	TRICLO	98	95	85	99	99	98	97
trimethoprim	TRIMET	99	85	80	99	99	99	94

5.4. RESULTS AND DISCUSSION

5.4.1 Oxidation and ARR hybrid

Figure 5.3a illustrates the removal of OMPs, encompassing a range of compound properties, by ARR and oxidation (O_3, AOP, and UV-AOP). The OMPs and their abbreviations are shown in *Table 5.2*. Quadrant I comprises OMPs which are poorly removed by both ARR and oxidation: there are a few compounds which are resistant to the processes and TCEP, a chlorinated aliphatic compound and a flame retardant, is an example. None of the compounds appear in quadrant II, corresponding to better removal by ARR and poor removal by oxidation, which is consistent with the relative effectiveness of chemical versus (micro) biological oxidation. Nevertheless, the use of this hybrid system is also to eliminate the metabolites/partial oxidation products produced due to oxidation. Quadrant III indicates that several OMPs are well-removed by both ARR and oxidation although iopromide, meprobamate and N, N - diethyl - 3-methylbenzamide appear to be better-removed by ARR than oxidation. Carbamazepine and dilantin are recalcitrant towards ARR but well-removed by oxidation as seen in quadrant IV. The individual quadrant plots between ARR - ozone, ARR - AOP and ARR - UV-AOP are shown in *Figures 5.3b, 5.3c, 5.3d,* respectively.

O_3 can be used in place of AOP in the hybrid process, although the percent removals of a few compounds decrease significantly. However, an attribute of O_3 is that it is a better disinfectant and can act as another pathogen barrier. A potential process constraint is bromate formation, but any bromate formed can potentially be eliminated by ARR under anoxic conditions [17]. Another constraint is the high oxidant demand of the feed-water; the alternative hybrid of ARR as a pre-treatment to O_3 would result in a lower oxidant demand, but the role of ARR as a metabolite barrier would be lost. UV-AOP as a post-treatment to ARR would also disinfect the water and, since it involves only $^{\bullet}$OH radicals, there is no bromate formation. As seen, carbamazepine and dilantin are poorly removed by ARR but well-removed by UV-AOP, while TCEP is recalcitrant to both UV-AOP and ARR.

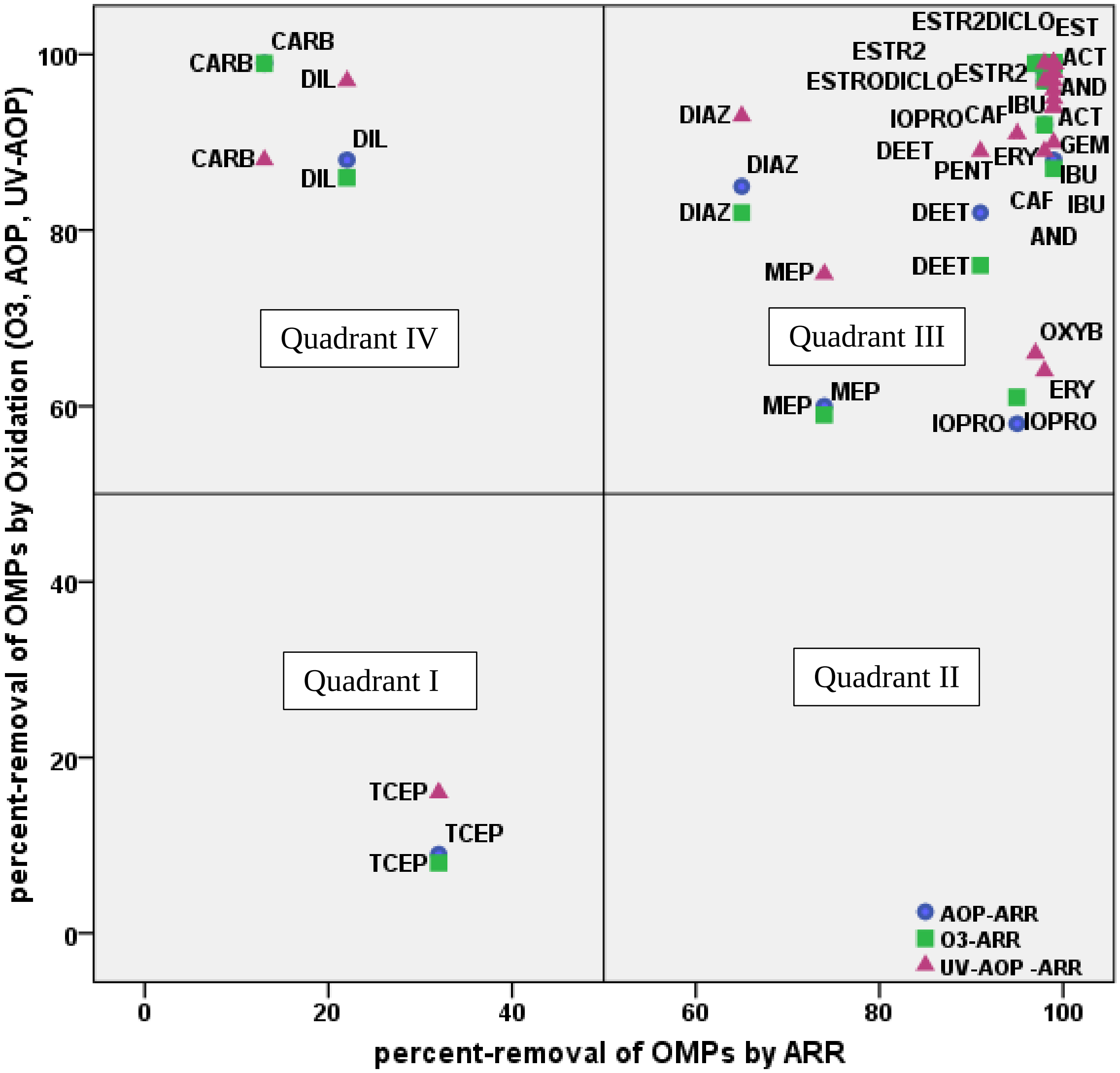


Figure 5.3a: OMPs attenuation by ARR and oxidation (O_3, AOP and UV-AOP).
The blue circles indicate AOP-ARR; the green squares indicate O3-ARR; and maroon triangles indicate UV-AOP-ARR hybrids

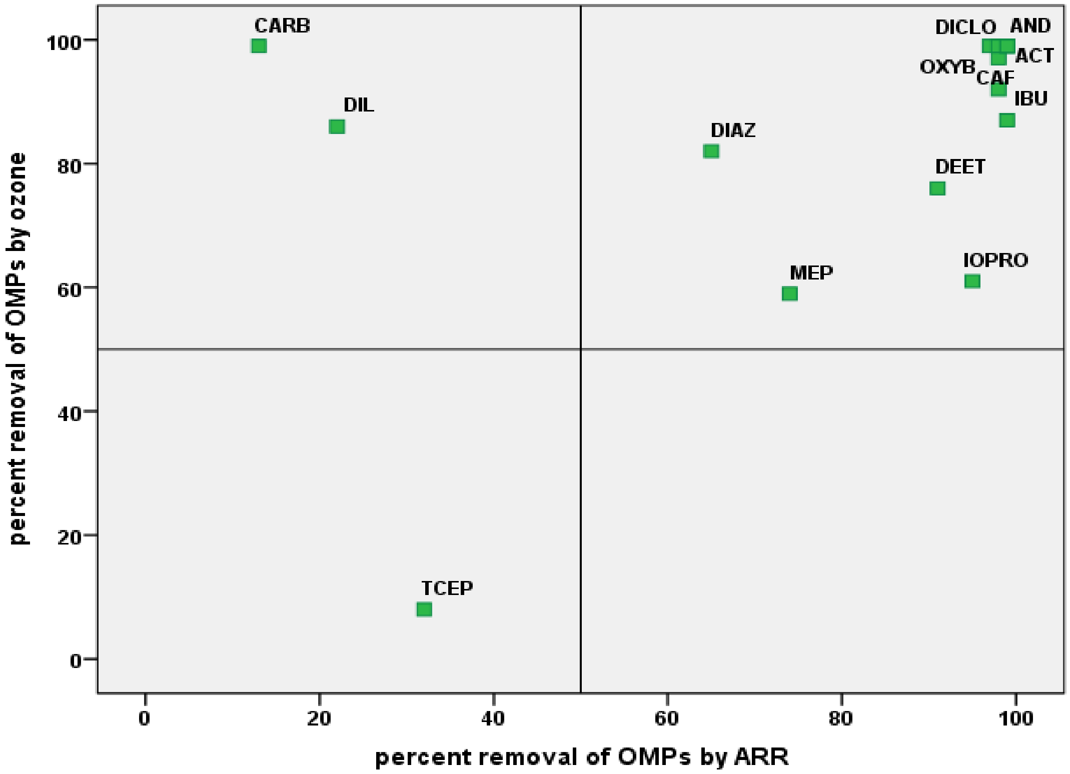


Figure 5.3b: OMPs attenuation by ARR and ozone

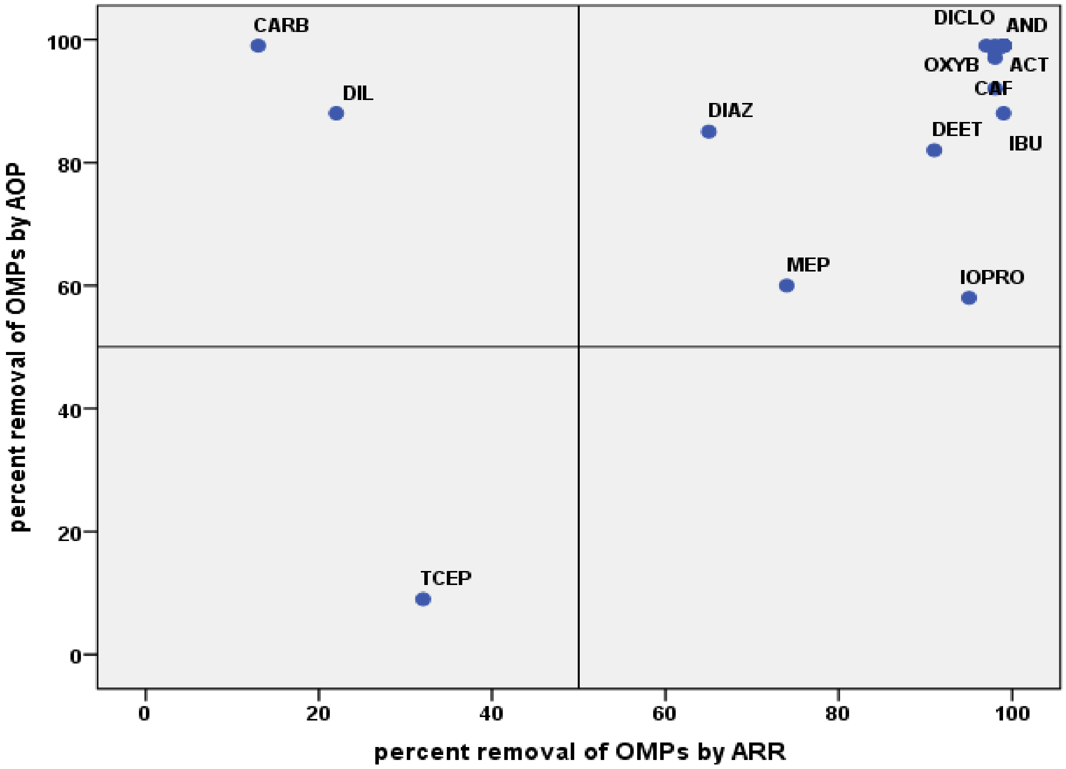


Figure 5.3c: OMPs attenuation by ARR vs advanced oxidation process (AOP)

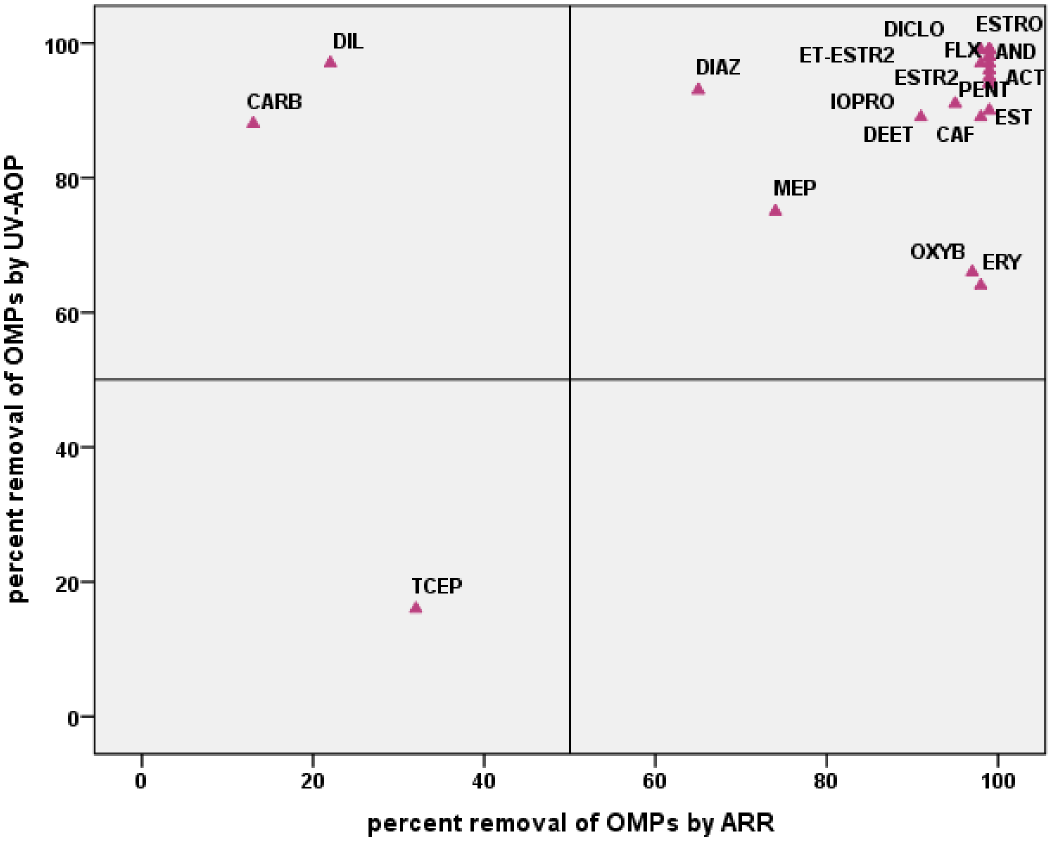


Figure 5.3d: OMPs attenuation by ARR vs UV-AOP

5.4.2 Adsorption and ARR hybrid

Figure 5.4 illustrates the percent removals of a number of OMPs, encompassing a range of compound properties, by ARR and GAC. TCEP is poorly removed by both processes whereas carbamazepine is better removed by GAC than ARR as seen in quadrant I and IV, respectively. Compounds such as naproxen, ibuprofen and gemfibrozil are better-removed by ARR than GAC as seen in quadrant II. Several OMPs present in quadrant III indicate good-removal by both GAC and ARR. The benefit of using ARR prior to GAC helps to reduce the DOC loading onto the GAC

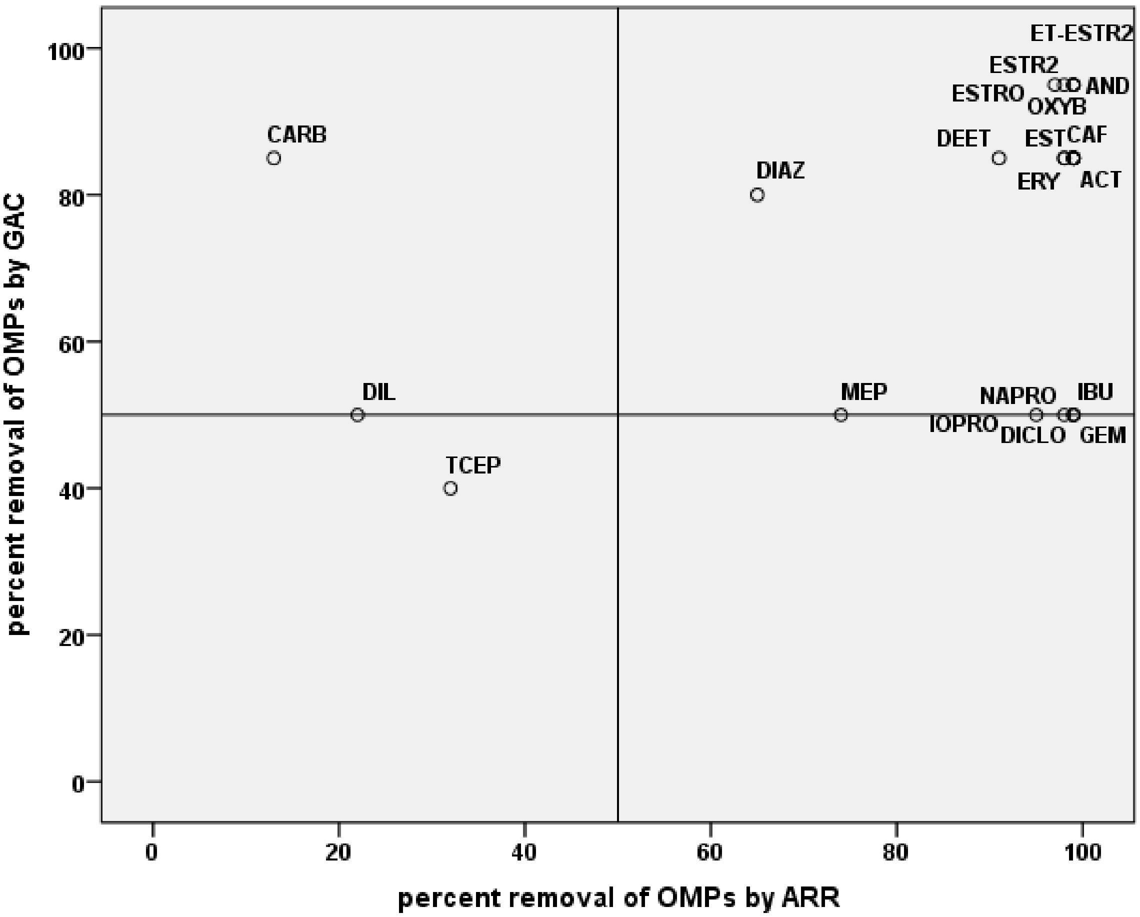


Figure 5.4: OMPs attenuation by ARR and GAC

5.4.3 Membrane and ARR hybrid

The synergy of the ARR-NF/RO, wherein ARR functions as a pre-treatment to NF or RO, benefits by minimizing membrane fouling [5]. Since ARR functions as a biofiltration process, organic foulants (e.g., proteins and polysaccharides) and biofoulants (e.g., carboxylic acids and aldehydes) can be effectively eliminated, particularly over ARR residence times of days to months [5]. In addition to being effective barriers for OMPs, both processes also provide effective barriers for pathogens. The industry-standard for non-potable reuse is using ARR after advanced treatment including RO. However, there are two arguments against this approach. First, NF is a lower-pressure (and hence lower cost) alternative to RO which provides greater OMP selectivity over background salts.

Studies have shown that NF is similar in efficiency to RO with respect to OMP removal and also the ARR/NF hybrid is very effective in pathogen and OMP removals from water [18-20]. Second, the use of RO/NF before ARR does not acknowledge the (pre) treatment attributes of ARR and instead relegates the ARR technology to only storage. Proponents of the industry standard argue that, without RO, one risks groundwater contamination. But the counterargument is that, with proper well construction and operation, one can dedicate and isolate part of the groundwater aquifer as a treatment zone [21-24]. Also, NF can be incorporated as a second microbial (virus) barrier. *Figure 5.5a* shows the OMP elimination by ARR and NF/RO. The absence of OMPs in quadrant I indicates that none of the OMPs were recalcitrant to both ARR and NF/RO. Acetaminophen and naproxen were better removed by ARR than NF as observed in quadrant II. A majority of the OMPs were well-removed by both ARR and NF as seen in quadrant III. TCEP, dilantin and carbamazepine, recalcitrant to ARR, were well-removed by NF/RO as seen in quadrant IV. In the case of RO, all the OMPs were very well-removed with above 85% removal efficiencies. However, the removal trends of OMPs for NF and RO were generally similar. The individual quadrant plots of ARR-RO and ARR-NF are shown in *Figure5.5b*, *5.5c*, respectively.

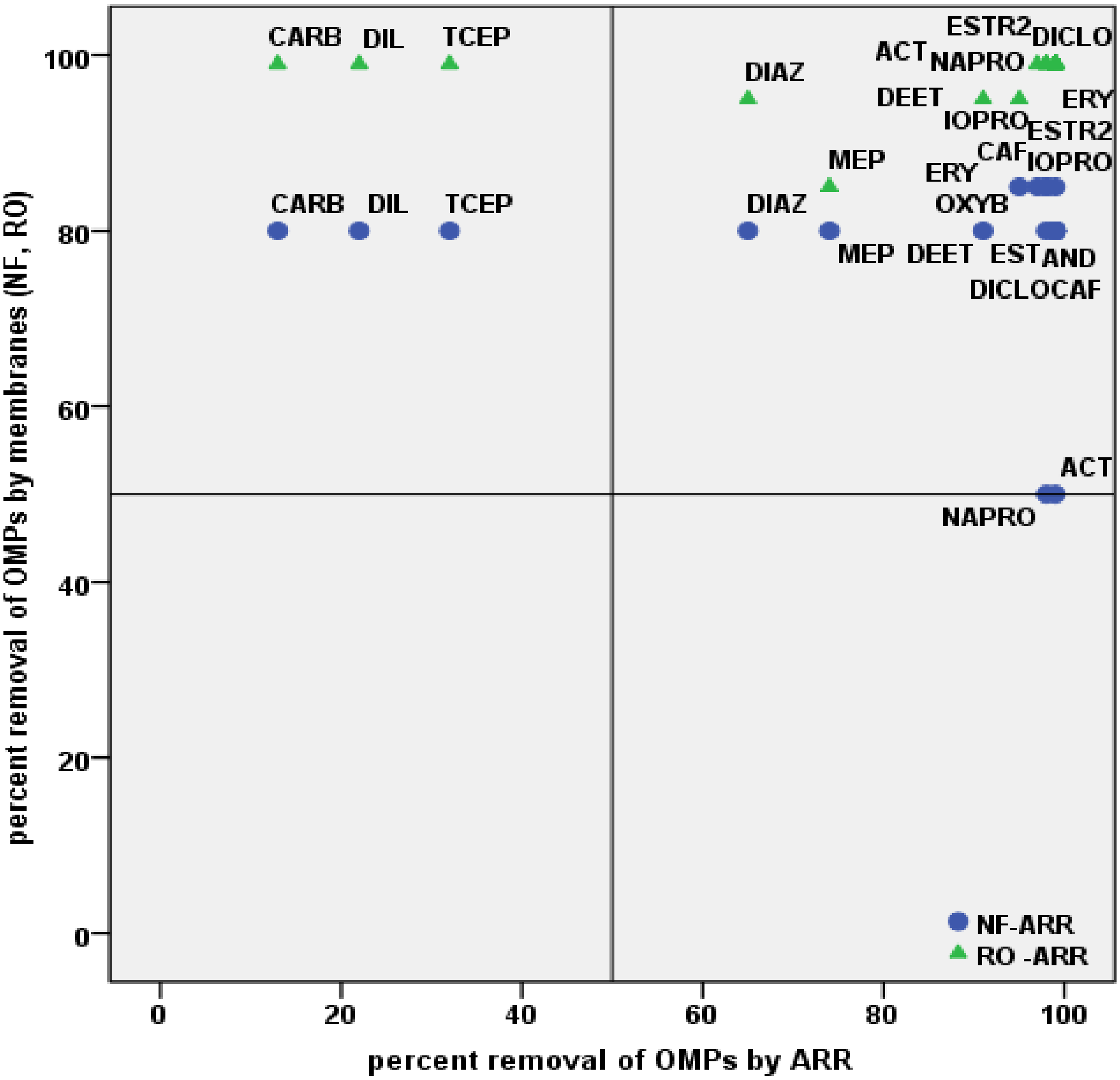


Figure 5.5a: OMPs attenuation by ARR and membranes (NF, RO).
The blue circles indicate NF-ARR hybrid and green triangles indicate RO-ARR hybrid

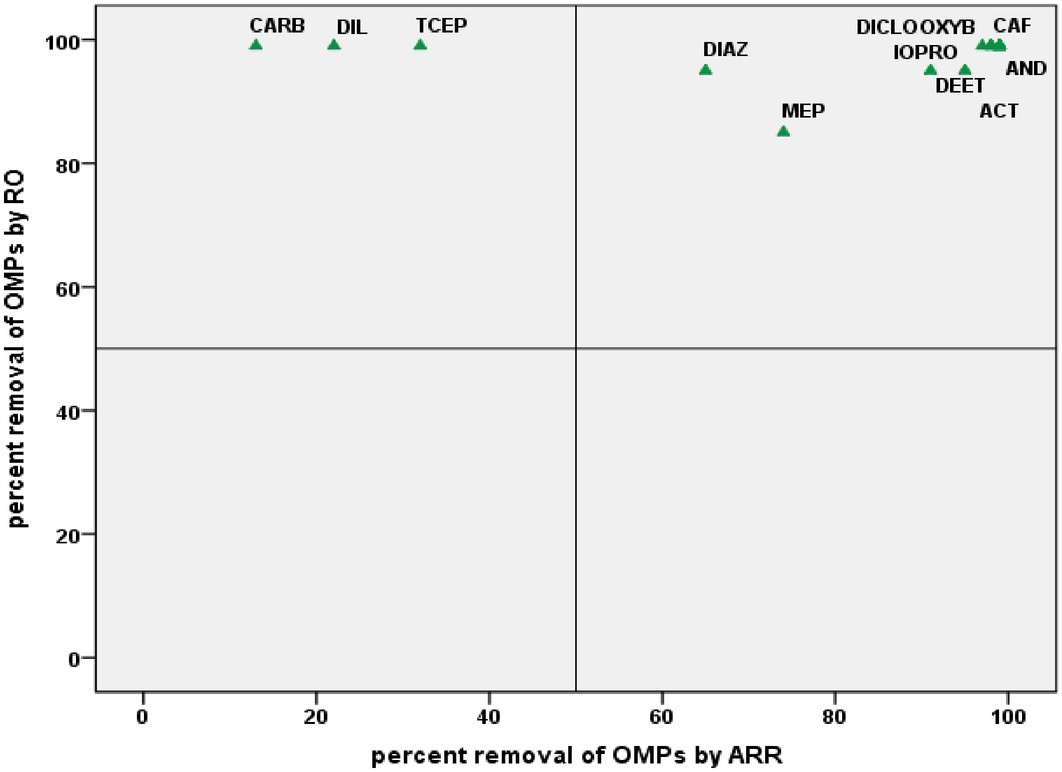


Figure 5.5b: OMPs attenuation by ARR and RO

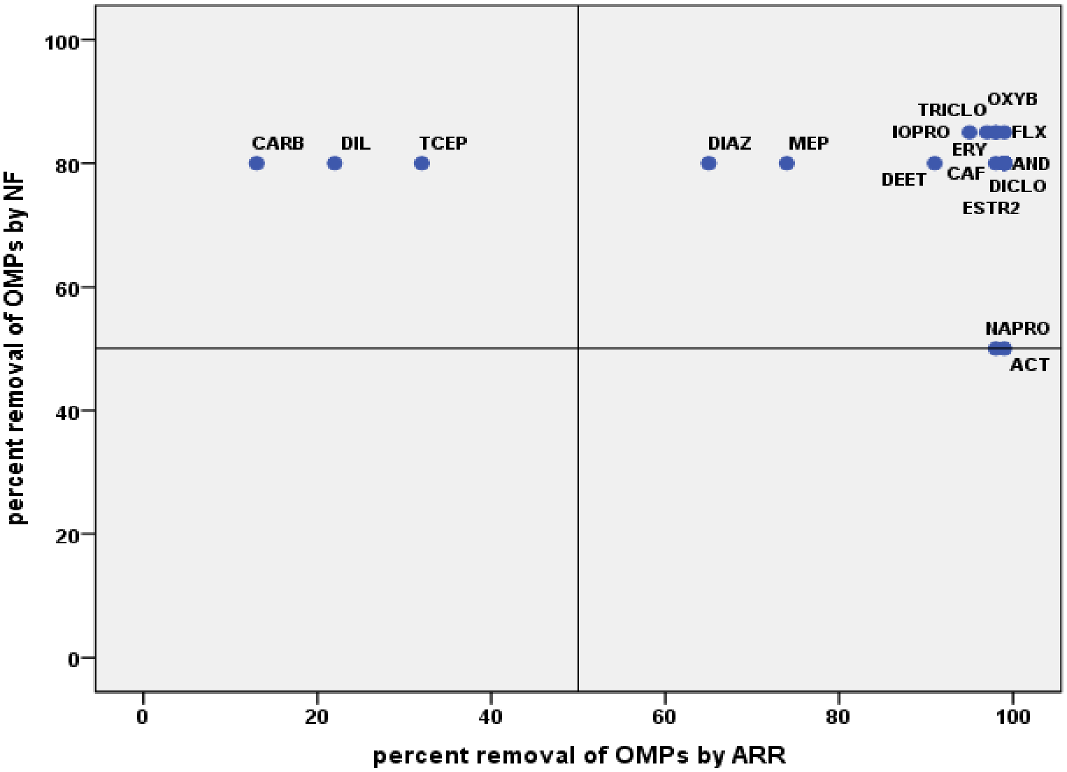


Figure 5.5c: OMPs attenuation by ARR and NF

5.5. CONCLUSION

As shown, the hybrids of processes help to establish a sustainable approach and compensate for the limitations of the individual processes. However, when oxidation by ozone is used as a treatment process, additional studies are required on the metabolites due to oxidation of OMPs [25]. Additionally, a cost analysis associated with operating the hybrids should be done, although the individual costs in operating the treatment plants are available in literature [26, 27]. Also, it would be interesting to study the performance amongst the advanced treatment processes for the OMPs removal thereby creating a better multi-barrier approach for sustainable and efficient removal of OMPs.

REFERENCES

1. Dordio, A., Carvalho, A.J.P., Teixeira, D.M., Dias, C.B., Pinto, A.P., *Removal of pharmaceuticals in microcosm constructed wetlands using Typha spp. and LECA.* Bioresource Technology, 2010. **101**(3): p. 886-892.

2. Snyder, S.A., Wert, E.C., Lei, H., Westerhoff, P., Yoon, Y., *Removal of EDCs and Pharmaceuticals in Drinking and Reuse Treatment Processes. Project # [2758]*. 2007, Awwa Research Foundation Denver, Colorado. p. 331.

3. Aharoni, A., Guttman, J., Cikurel, H., *Guidelines for Design, Operation and Maintenance of SAT (and hybrid SAT) Systems*. 2011, Mekorot and UNESCO-IHE; 018530-SWITCH.

4. Kazner, C., Baghoth, S., Sharma, S., Amy, G., Wintgens, T., Melin, T., *Comparing the effluent organic matter removal of direct NF and powdered activated carbon/NF as high quality pretreatment options for artificial groundwater recharge.* Water Science and Technology, 2008. **57**(6): p. 821-827.

5. Cikurel, H., Sharma, S.K., Jekel, M., Kazner, C., Wintgens, T., Amy, G., Ernst, M., Guttman, Y., Tal, N., Kreitzer, T., Putschew, A., Vairavamoorthy, K., Aharoni, A. *Alternative hybrid SAT-membrane treatments: Short SAT-NF treatment to upgrade effluent quality*. in *Proceedings of the IWA Regional Conference on Membrane Technology and Water Reuse*. 2010. Istanbul, Turkey.

6. Ray, C., Grischek, T., Schubert, J., Wang, J.Z., Speth, T.F., *A perspective of riverbank filtration.* Journal- American Water Works Association, 2002. **94**(4): p. 149-160.

7. Wang, J.Z., Hubbs, S.A., Song, R., *Evaluation of Riverbank Filtration as a Drinking Water Treatment Process*. 2002, Denver: AWWA Research Foundation and American Water Works Association.

8. Berger, P., *Removal of Cryptosporidium using bank filtration* in *Riverbank Filtration: Understanding Contaminant Biogeochemistry and Pathogen Removal*, C. Ray, A.C.M. Bourg, and F. Laszlo, Editors. 2003, Kluwer Academic, Dordrecht, The Netherlands. p. 85-121.

9. Schijven, J., Berger, P., Miettinen, I., *Removal of pathogens, surrogates, indicators, and toxins using riverbank filtration*, in *Riverbank Filtration*, C. Ray, G. Melin, and R. Linsky, Editors. 2003, Springer Netherlands. p. 73-116.

10. von Gunten, U., *Ozonation of drinking water: Part I. Oxidation kinetics and product formation.* Water Research, 2003. **37**(7): p. 1443-1467.

11. Moreno-Castilla, C., *Adsorption of organic molecules from aqueous solutions on carbon materials.* Carbon, 2004. **42**(1): p. 83-94.

12. Dobbs, R.A.,Cohen, J.M., *Carbon adsorption isotherms for toxic organics, Municipal environmental research laboratory, Office of research and development, US Environmental Protection Agency*, in *Ohio. Report No. EPA-60018-80-023*. 1980.

13. Villacañas, F., Pereira, M.F.R., Órfão, J.J.M., Figueiredo, J.L., *Adsorption of simple aromatic compounds on activated carbons.* Journal of Colloid and Interface Science, 2006. **293**(1): p. 128-136.

14. Chen, W., Duan, L., Zhu, D., *Adsorption of polar and nonpolar organic chemicals to carbon nanotubes.* Environmental Science and Technology, 2007. **41**(24): p. 8295-8300.

15. El-Dib, M.A.,Badawy, M.I., *Adsorption of soluble aromatic hydrocarbons on granular activated carbon.* Water Research, 1979. **13**(3): p. 255-258.

16. Quiñones, I., C Ford, J., Guiochon, G., *Multisolute adsorption equilibria in a reversed-phase liquid chromatography system.* Chemical Engineering Science, 2000. **55**(5): p. 909-929.

17. Hijnen, W.A.M., Jong, R., van der Kooij, D., *Bromate removal in a denitrifying bioreactor used in water treatment.* Water Research, 1999. **33**(4): p. 1049-1053.

18. Wintgens, T., Salehi, F., Hochstrat, R., Melin, T., *Emerging contaminants and treatment options in water recycling for indirect potable use.* Water Science and Technology, 2008. **57**(1): p. 99-108.

19. Bellona, C., Drewes, J.E., Xu, P., Amy, G., *Factors affecting the rejection of organic solutes during NF/RO treatment--a literature review.* Water Research, 2004. **38**(12): p. 2795-2809.

20. Aguilar, A., Jimenez, B., Becerril, J., Castro, L., *Use of nanofiltration for potable water from an aquifer recharged with wastewater.* Water Science and Technology, 2008. **57**(6): p. 927.

21. Vymazal, J., *Constructed wetlands for wastewater treatment: Five decades of experience.* Environmental Science and Technology, 2010. **45**(1): p. 61-69.

22. Budd, R., O'Geen, A., Goh, K.S., Bondarenko, S., Gan, J., *Efficacy of constructed wetlands in pesticide removal from tailwaters in the central valley, California.* Environmental Science and Technology, 2009. **43**(8): p. 2925-2930.

23. Knight, R.L., Kadlec, R.H., Ohlendorf, H.M., *The use of treatment wetlands for petroleum industry effluents.* Environmental Science and Technology, 1999. **33**(7): p. 973-980.

24. Cole, S., *The Emergence of Treatment Wetlands.* Environmental Science and Technology, 1998. **32**(9): p. 218A-223A.

25. Mawhinney, D.B., Vanderford, B.J., Snyder, S.A., *Transformation of 1H-benzotriazole by ozone in aqueous solution.* Environmental Science and Technology, 2012. **46**(13): p. 7102-7111.

26. Lienert, J., Koller, M., Konrad, J., McArdell, C.S., Schuwirth, N., *Multiple-criteria decision analysis reveals high stakeholder preference to remove pharmaceuticals from hospital wastewater.* Environmental Science and Technology, 2011. **45**(9): p. 3848-3857.

27. Sudhakaran, S., Lattemann, S., Amy, G.L., *Appropriate drinking water treatment processes for organic micropollutants removal based on experimental and model studies — A multi-criteria analysis study*. Science of Total Environment, 2013. **442**(0): p. 478-488.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

- QSAR models based on multi-linear regression and artificial neural networks were developed for OMPs removal; several models were developed based on multi-linear regression so that the constructed models could be updated for future use and easily reproduced for further learning and understanding.
- New chemical descriptors such as the energy gap between the orbitals (E_{LUMO}-E_{HOMO}), ionisation potential, electron affinity, mean oxidation number, and halogen surface area, which were integrated in the QSAR models, helped to understand ozonation (oxidation) in a more comprehensive manner.
- QSAR models were developed for both the performance indices of oxidation (percent-removal of OMPs and ozone and hydroxyl radical rate constants (k_{O3}, k_{OH})). It was concluded that rate constants were a better index for QSAR modeling purposes compared to percent removals since they were influenced by lesser boundary conditions. Also, the rate constants could be coupled with ideal chemical reactors to predict process performance.
- A DSS based on MCA was developed for experimental and QSAR model studies. RBF and oxidation were the preferred processes for OMPs removal. The DSS can be helpful in the experimental set-up and planning of drinking water treatment trains for OMPs removal. However, the rankings obtained in this study are valid for the given set of alternatives, criteria, scores, weight and can be revised/refined, if new or better data are available.

- Hybrids of natural systems with advanced treatment processes help to develop a sustainable and efficient approach for OMPs removal compensating for the limitations of the individual processes.

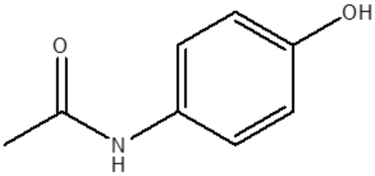
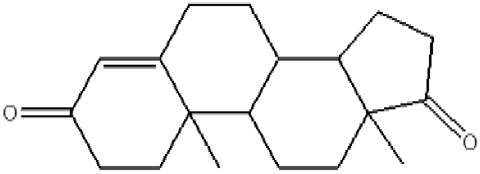
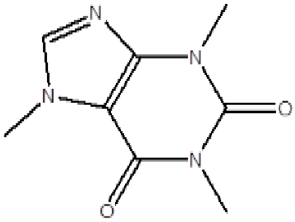
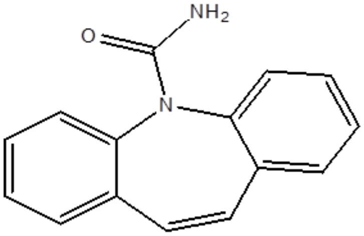
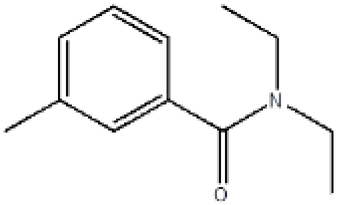
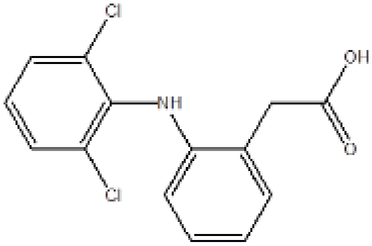
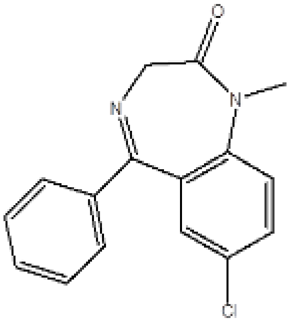
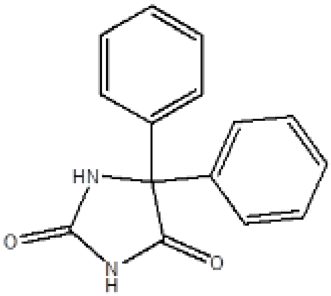
6.2. RECOMMENDATIONS

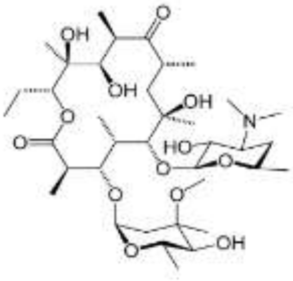
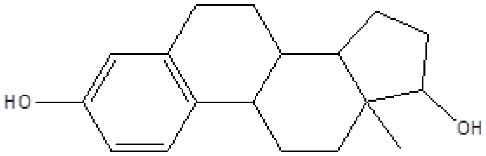
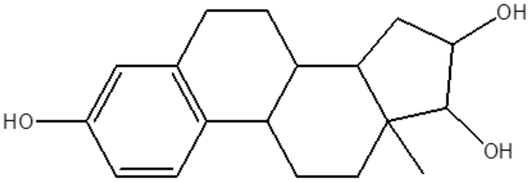
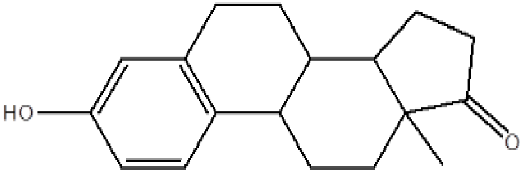
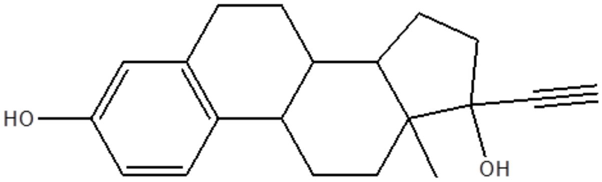
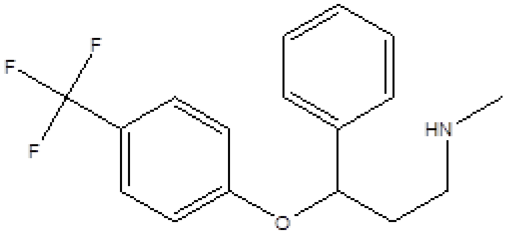
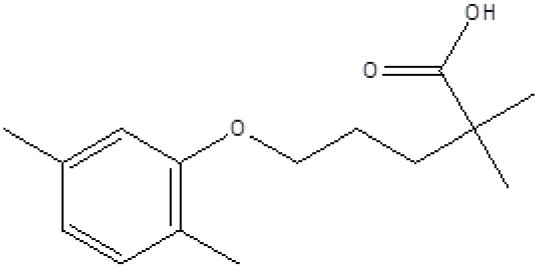
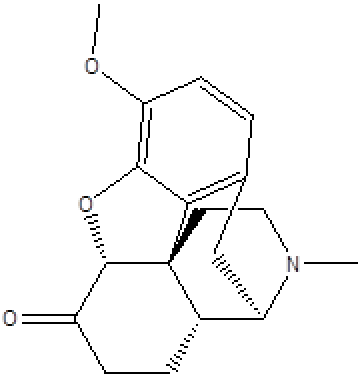
- The different QSAR models built for oxidation (ozonation/AOP) in *Chapter 2* and *3* and others available in the scientific literature could be compiled and developed into a software program which can compute the oxidation efficiency of the OMPs
- QSAR models for the oxidation metabolites of OMPs should be the focus of further research since studies have shown their negative impacts
- The research studies in *Chapter 2* have shown strong correlation between energy gap of the molecular orbitals ($E_{LUMO} – E_{HOMO}$; energy of the lowest unoccupied molecular orbital and highest occupied molecular orbitals) and ozonation capacity. Future research should focus on developing a comprehensive understanding about the molecular orbitals of OMPs and ozone interactions.
- In the DSS developed in *Chapter 4*, additional percent removal rates for a wider range of OMPs for different water treatment processes can be incorporated based on QSAR modeling to expand the DSS to a wider decision-making context.
- Based on the results obtained for the hybrid systems in *Chapter 5*, a cost analysis (both capital and operation/maintenance) in operating the hybrids should be performed to promote this approach. Also, for an efficient multi-barrier approach for OMPs removal, studies should be focused on optimal/sustainable processes (e.g., NF be replaced by RO) involved in the water treatment train.

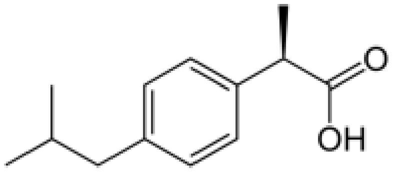
APPENDICES

APPENDIX A

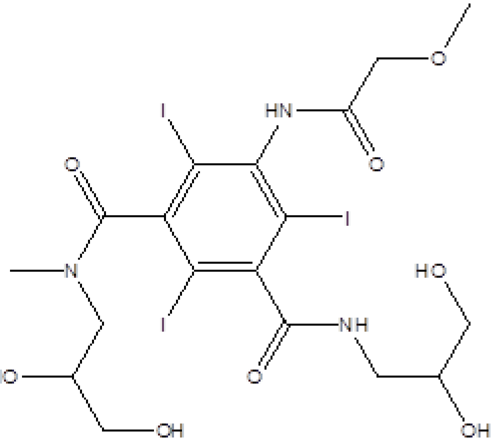
A1: Chemical Structures of OMPs involved in *Chapters 2 and 5*

 Acetaminophen	 Androstenedione
 Caffeine	 Carbamzepine
 DEET	 Diclofenac
 Diazepam	 Dilantin

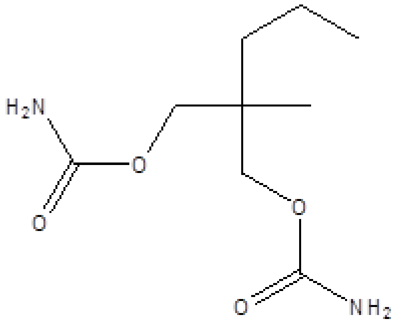
 Erythromycin	 Estradiol
 Estriol	 Estrone
 Ethinyl Estradiol	 Fluoxetine
 Gemfibrozil	 Hydrocodone



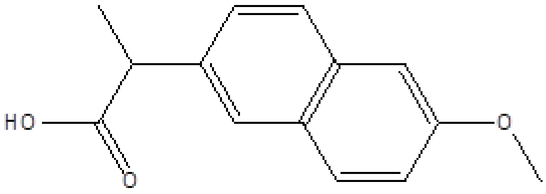
Ibuprofen



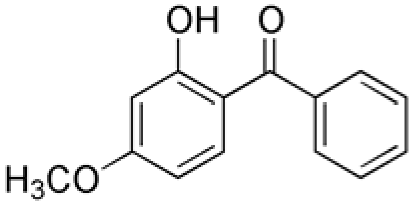
Iopromide



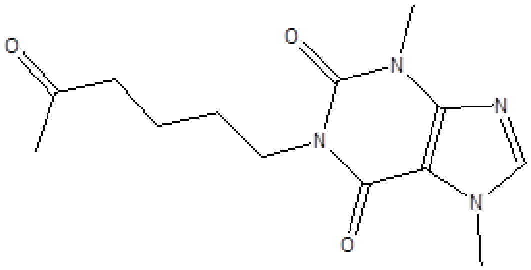
Meprobamate



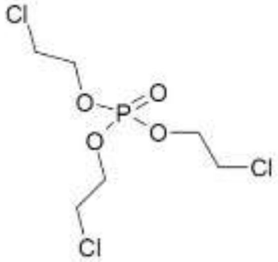
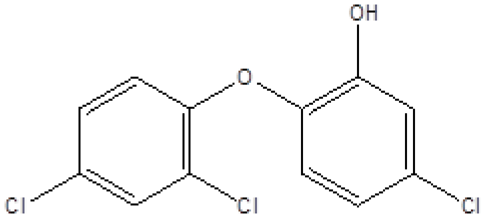
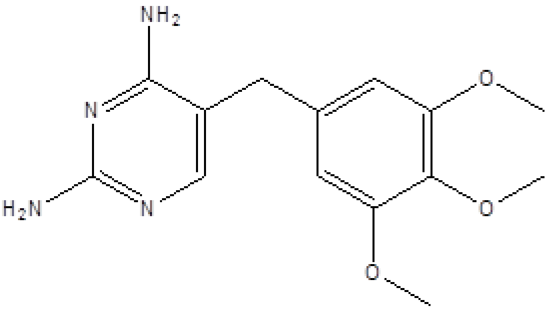
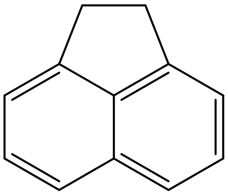
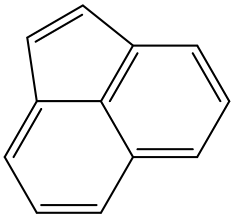
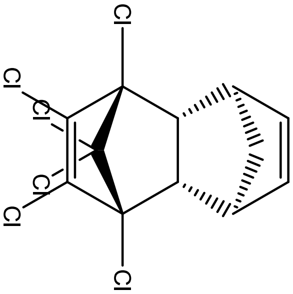
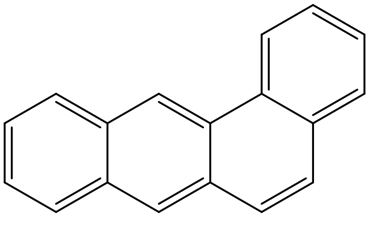
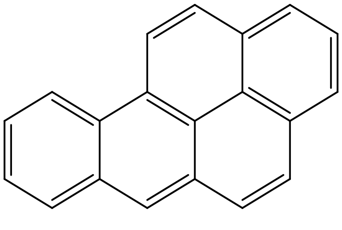
Naproxen

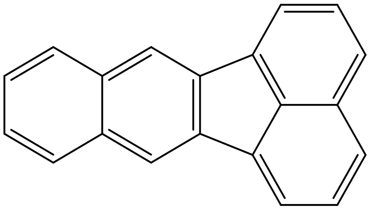
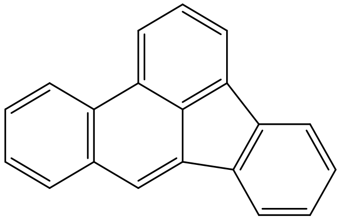
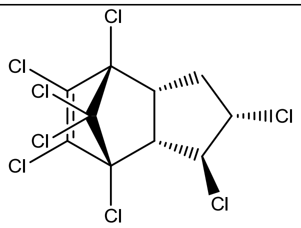
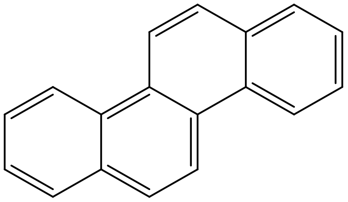
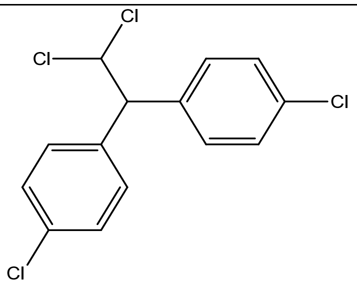
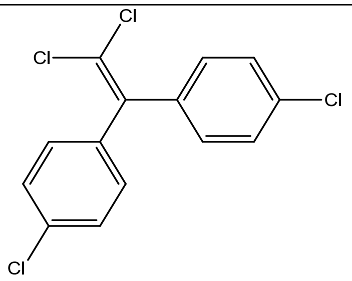
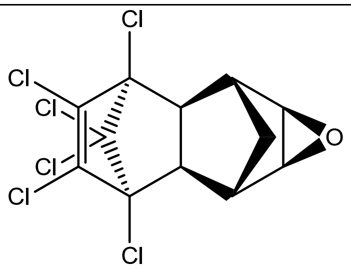
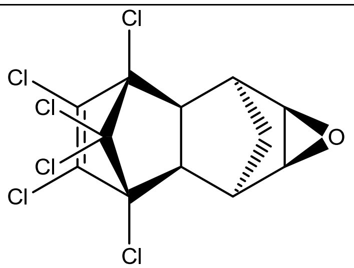
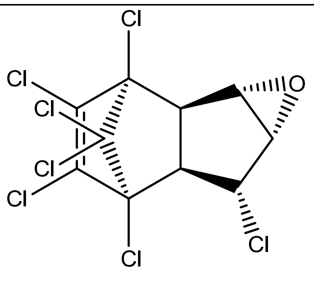
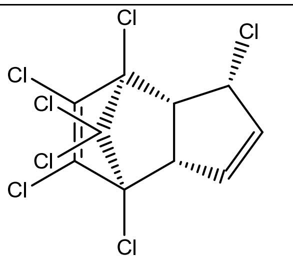


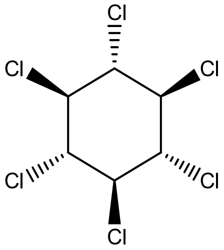
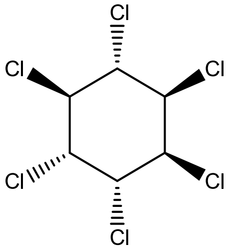
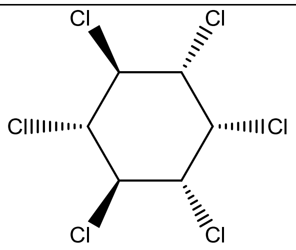
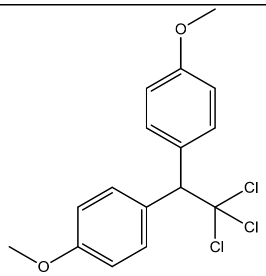
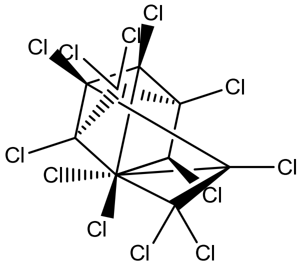
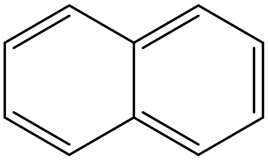
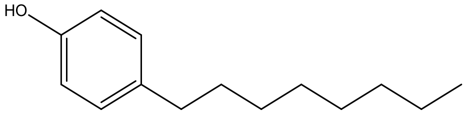
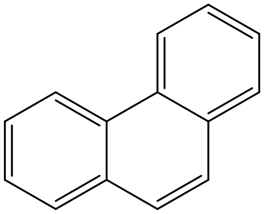
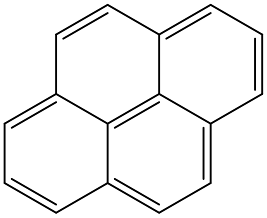
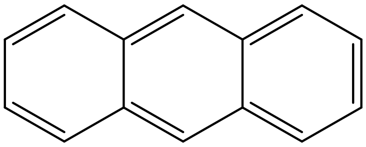
Oxybenzone



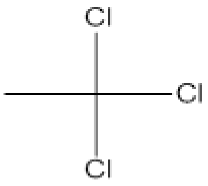
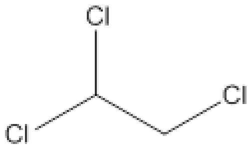
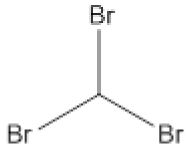
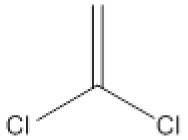
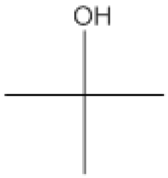
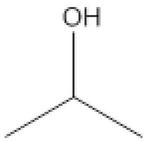
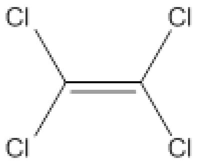
Pentoxyfylline

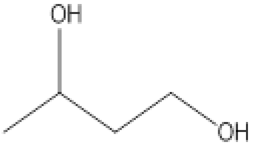
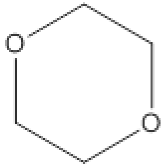
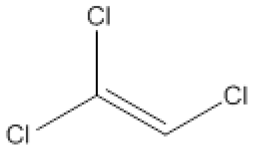
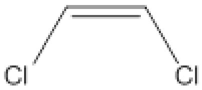
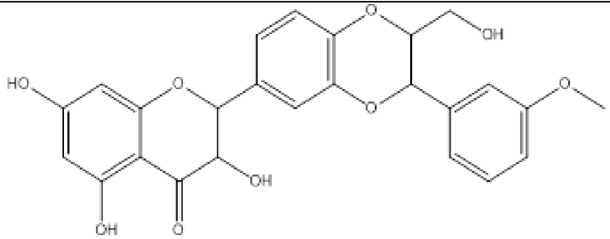
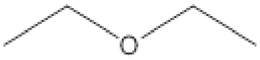
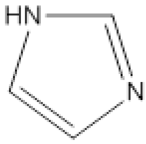
 TCEP	 Triclosan
 Trimethoprim	 acenaphthene
 acenaphthylene	 aldrin
 Benzo(a)anthracene	 Benzo(a)pyrene

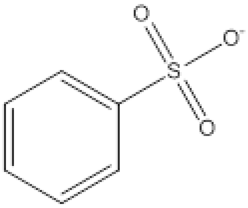
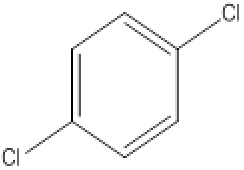
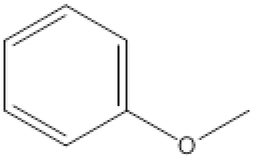
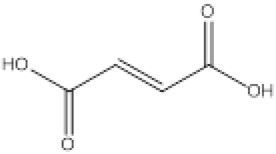
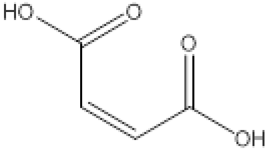
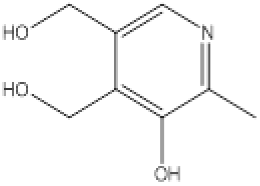
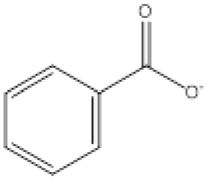
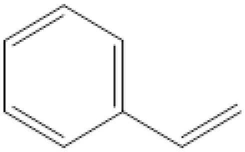
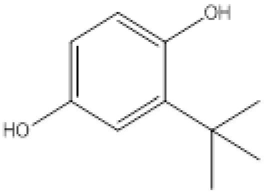
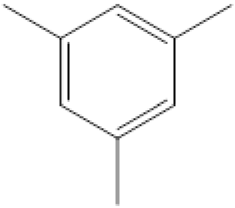
 Benzo(k)fluoranthene	 benzo(b)fluoranthene
 chlordane	 chrysene
 DDD	 DDE
 Dieldrin	 endrin
 Heptachlor epoxide	 heptachlor

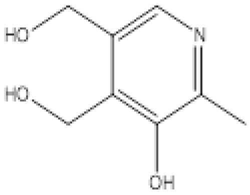
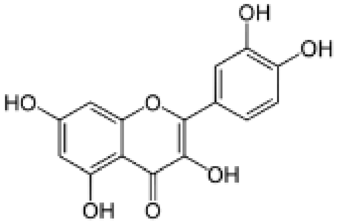
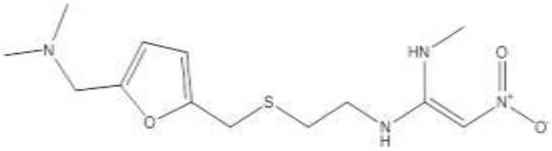
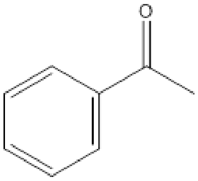
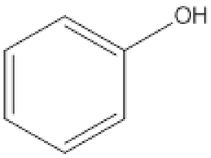
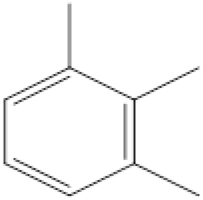
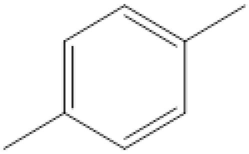
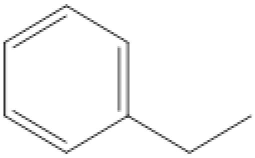
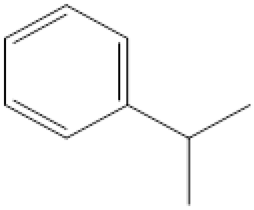
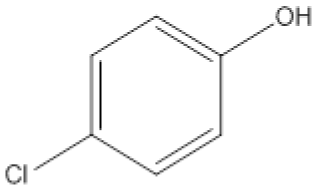
 β-lindane	 α-lindane
 δ-lindane	 methoxychlor
 mirex	 napthalene
 octylphenol	 phenanthrene
 pyrene	 anthracene

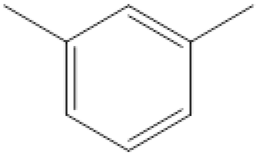
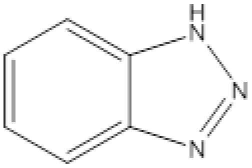
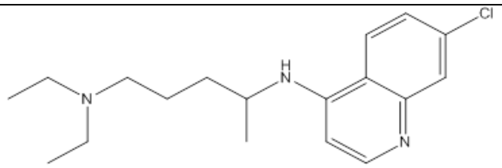
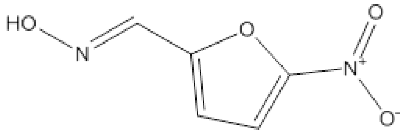
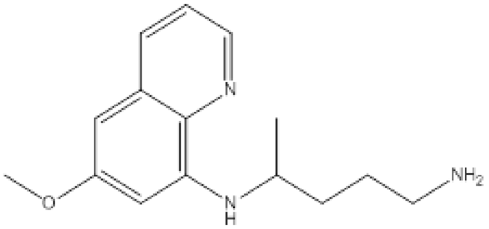
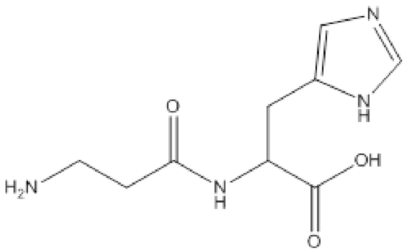
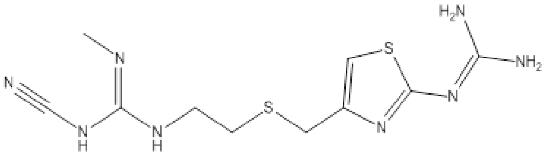
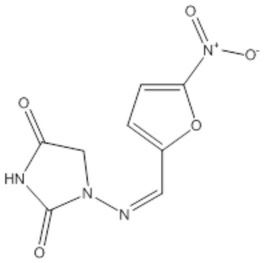
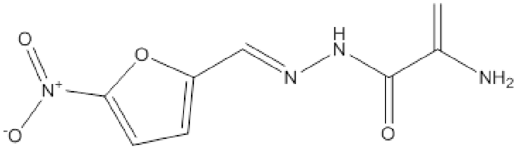
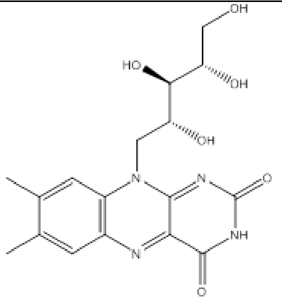
A2: Chemical Structures of OMPs involved in *Chapter 3*

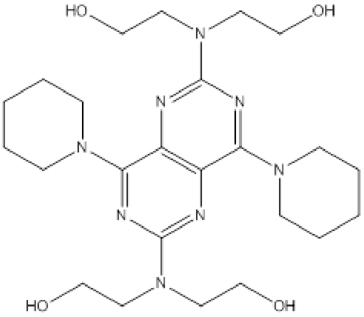
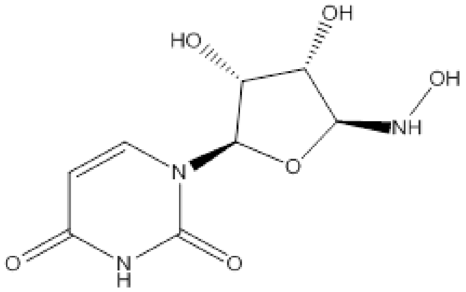
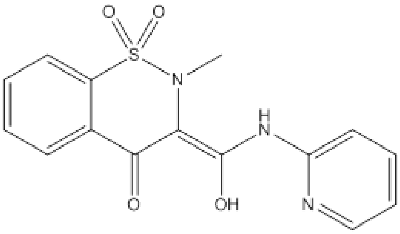
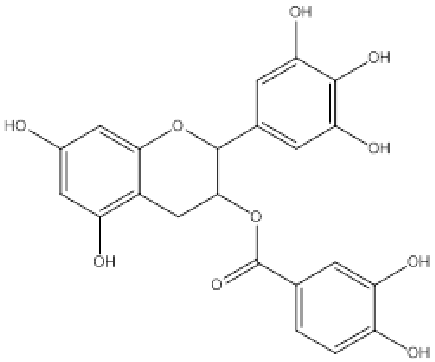
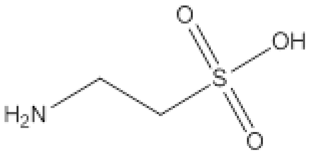
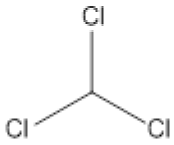
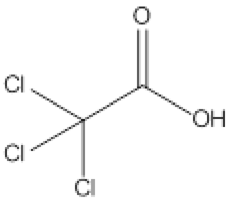
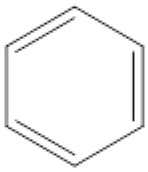
 1,1,1-trichloroethane	 dibromomethane
 1,1,2-trichloroethane	 bromoform
 1,1-dichloroethene	 1,2-dibromoethane
 tert-butanol	 methanol
 propan-2-ol	 tetrachloroethene

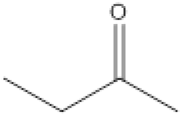
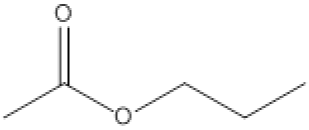
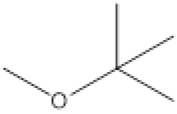
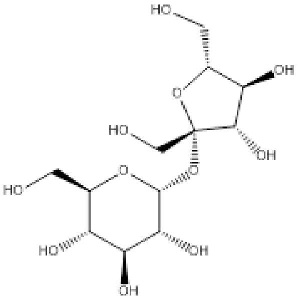
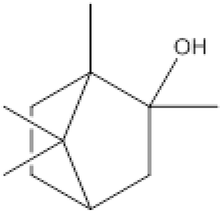
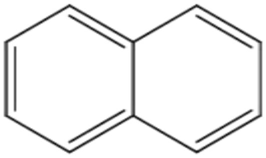
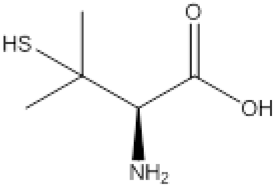
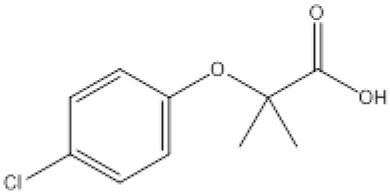
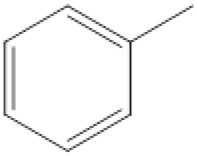
 butan-1,3-diol	 ethanol
 dioxane	 propanol
 trichloroethene	 Cis-dichloroethene
 silybin	 diethylether
 imidazole	 butanol

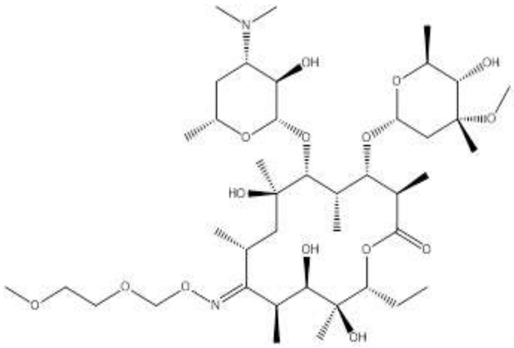
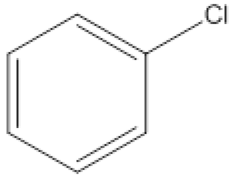
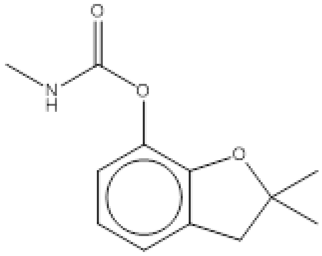
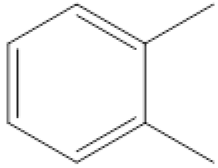
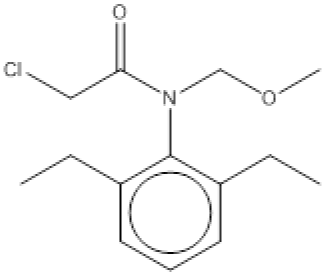
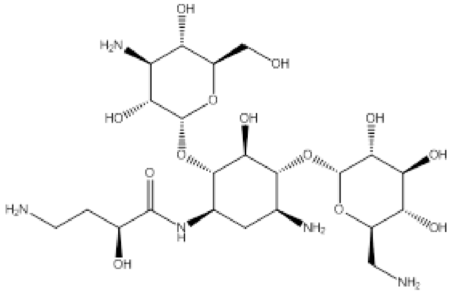
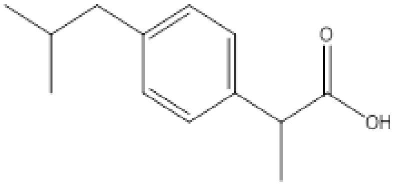
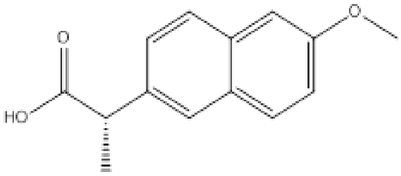
 benzene sulfonate	 1,4-dichlorobenzene
 anisole	 fumaric acid
 maleic acid	 pyridoxine
 benzoate ion	 styrene
 tert-butylhydroquinone	 1,3,5-trimethylbenzene

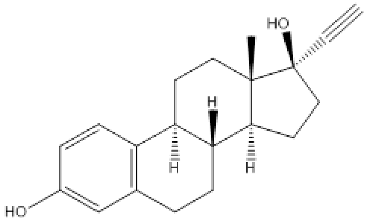
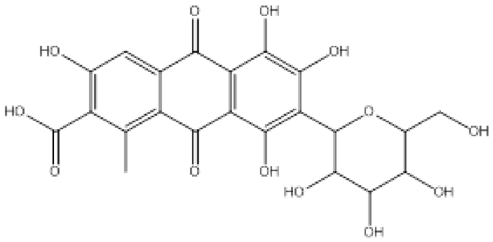
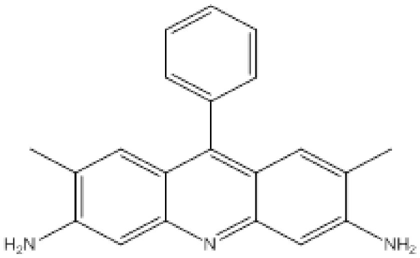
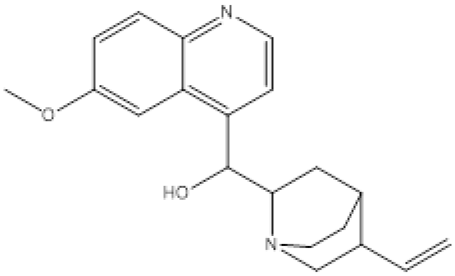
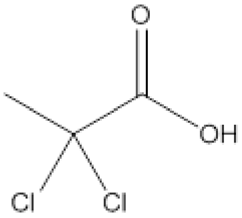
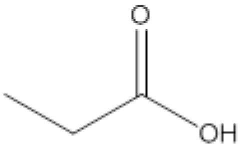
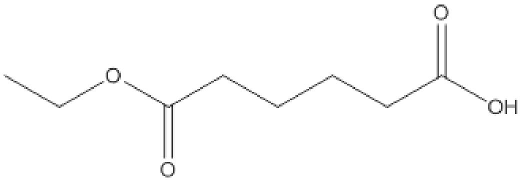
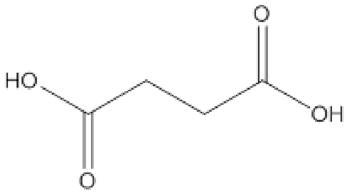
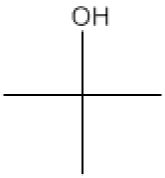
 pyridoxine	 quercetin
 ranitidine	 acetophenone
 phenol	 1,2,3-trimethylbenzene
 para-xylene	 ethyl benzene
 iso-propyl benzene	 4-chlorophenol

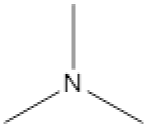
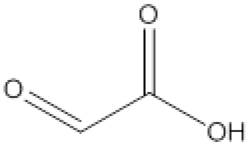
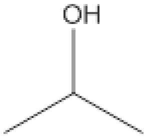
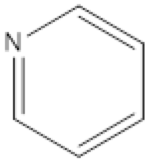
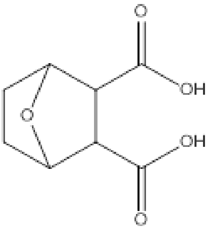
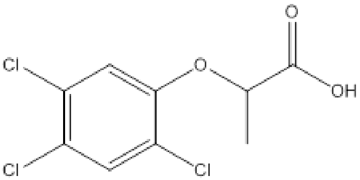
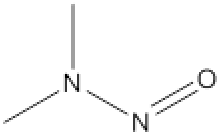
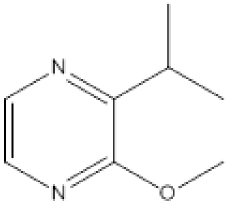
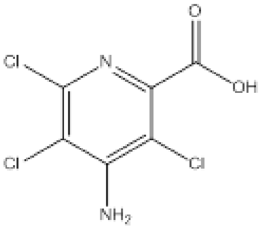
 meta-xylene	 benzotriazole
 chloroquine	 nifuroxime
 primaquine	 carnosine
 tiotidine	 nitrofurantoin
 nifuraldezone	 riboflavin

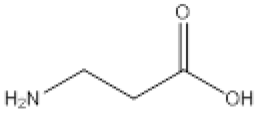
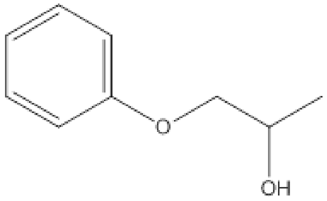
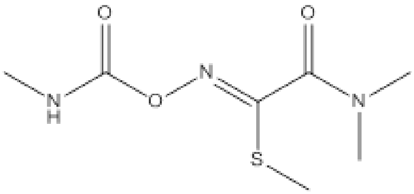
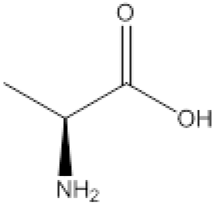
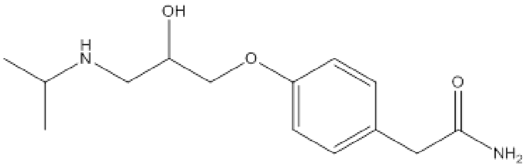
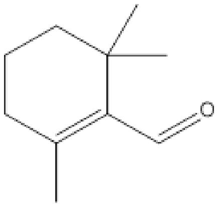
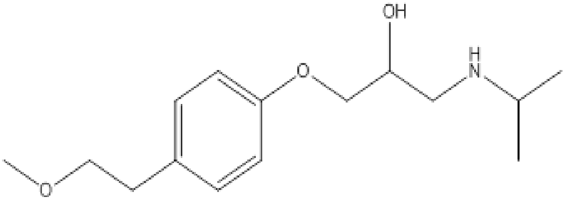
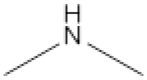
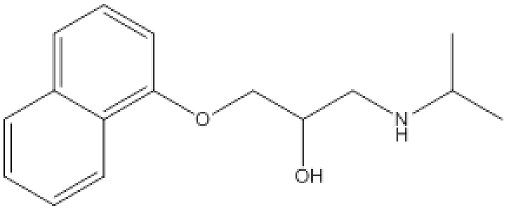
 dipyridamole	 azauridine
 piroxicam	 EGCG
 taurine	 chloroform
 trichloroacetic acid	 benzene
 spermidine	 formaldehyde

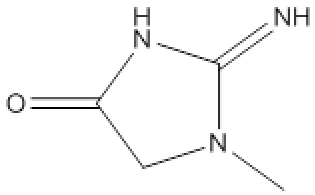
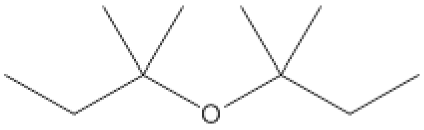
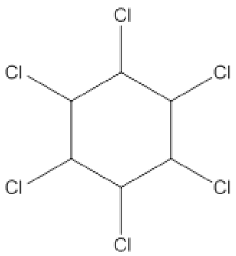
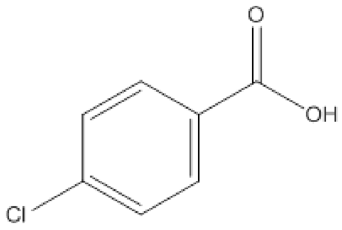
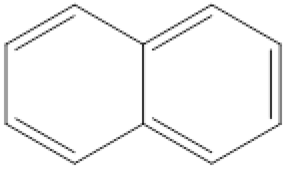
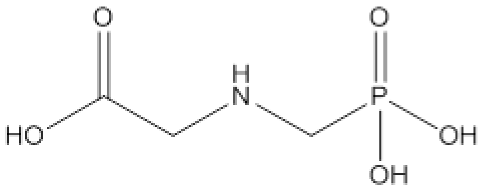
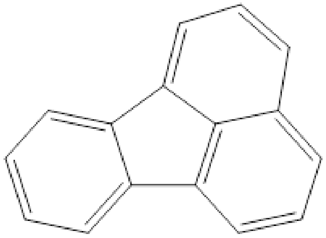
 propanal	 butan-2-one
 propyl acetate	 methyl tert-butyl ether (MTBE)
 sucrose	 2-methyl iso borneol
 napthalene	 pencillamine
 clofibric acid	 toluene

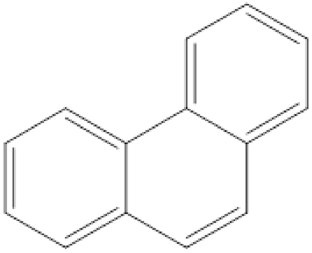
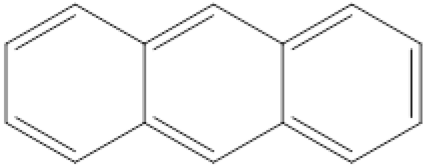
 roxithromycin	 chlorobenzene
 carbafuran	 o-xylene
 alachlor	 amikacin
 Ibuprofen	 naproxen

 17α-ethinyl estradiol	 carmine
 benzoflavine	 quinidine
 dalapon	 propionic acid
 adipic acid monoethyl ester	 succinic acid
 t-butanol	 methanol

 cyanic acid	 trimethylamine
 glyoxylic acid	 propan-2-ol
 pyridine	 octanal
 endothall	 fenoprop
 nitrosodimethylamine	 methylamine
 2-isopropyl-3-methoxypyrazine	 picloram

 β-alanine	 1-phenoxy-2-propanol
 oxamyl	 α-alanine
 atenolol	 β-cyclocitral
 metoprolol	 dimethylamine
 propanolol	 glutaric acid

 methanol	 creatinine
 MTBE	 lindane
 butylamine	 p-chlorobenzoic acid
 ethanol	 napthalene
 glyphosate	 fluoranthene

 phenanthrene	 anthracene

Appendix B

B1: Correlations between molecular descriptors and %removal of OMPs from *Chapter 2*

CRW-AOP	lnrmvl	CRW-O3	lnrml	ORW-AOP	%rml	ORW-O3	%rml
lnrmvl	1	lnrml(O3)	1	%rml(AOP)	1	%rml(O3)	1
X(Halo)	-.641**	X(Halo)	-.679**	X(Halo)	-.653**	X(Halo)	-.660**
WPSA	-.687**	WPSA	-.705**	WPSA	-.708**	WPSA	-.710**
volume	0.108	volume	0.113	volume	0.12	volume	0.08
SASA	0.067	SASA	0.08	SASA	0.085	SASA	0.036
QPpolrz	0.152	QPpolrz	0.152	QPpolrz	0.175	QPpolrz	0.143
PSA	0.089	PSA	0.098	PSA	0.163	PSA	0.114
PISA	.335*	PISA	.338*	PISA	0.336	PISA	.348*
O_C	-0.311	O_C	-0.216	O_C	-0.319	O_C	-0.343
N_C	-0.003	N_C	0.007	N_C	0.087	N_C	0.047
MON	-0.131	MON	-0.17	MON	-0.013	MON	-0.05
IP(eV)	-.554**	IP(eV)	-.541**	IP(eV)	-.544**	IP(eV)	-.550**
H_C	-0.209	H_C	-0.162	H_C	-0.221	H_C	-0.226
FOSA	0.106	FOSA	0.121	FOSA	0.099	FOSA	0.077
FISA	0.153	FISA	0.154	FISA	0.226	FISA	0.184
E_L-E_H	-.839**	E_L-E_H	-.895**	E_L-E_H	-.936**	E_L-E_H	-.902**
EA(eV)	-0.218	EA(eV)	-0.168	EA(eV)	-0.145	EA(eV)	-0.218
donorHB	0.145	donorHB	0.12	donorHB	0.202	donorHB	0.176
dipole	0.172	dipole	0.194	dipole	0.212	dipole	0.169
DBE	.542**	DBE	.548**	DBE	.561**	DBE	.537**
C=C	0.32	C=C	.334*	C=C	0.33	C=C	0.313
Arom	0.211	Arom	0.232	Arom	0.235	Arom	0.223
accptHB	0.058	accptHB	0.066	accptHB	0.084	accptHB	0.047
#rtvFG	-.592**	#rtvFG	-.646**	#rtvFG	-.600**	#rtvFG	-.537**
#rotor	0.066	#rotor	0.084	#rotor	0.043	#rotor	-0.008
#ringatoms	.451**	#ringatoms	.425**	#ringatoms	.456**	#ringatoms	.474**
#nonHatm	0.215	#nonHatm	0.218	#nonHatm	0.225	#nonHatm	0.186
#noncon	0.104	#noncon	0.065	#noncon	0.077	#noncon	0.114
#in56	.491**	#in56	.465**	#in56	.490**	#in56	.508**

**. Correlation is significant at the 0.01 level
*. Correlation is significant at the 0.05 level

B1: Correlations between molecular descriptors and %removal of OMPs from *Chapter 2*

PRW-AOP	lnrmvl	PRW-O3	lnrml	SRW-AOP	%rmvl	SRW-O3	rmvl
lnrmvl	1	lnrml	1	%rmvl	1	rmvl	1
X(Halo)	-.530**	X(Halo)	-.630**	X(Halo)	-.564**	X(Halo)	-.612**
WPSA	-.586**	WPSA	-.682**	WPSA	-.631**	WPSA	-.677**
volume	0.115	volume	0.148	volume	0.137	volume	0.157
SASA	0.071	SASA	0.107	SASA	0.083	SASA	0.097
QPpolrz	0.158	QPpolrz	0.191	QPpolrz	0.236	QPpolrz	0.235
PSA	0.077	PSA	0.152	PSA	0.161	PSA	0.185
PISA	.339*	PISA	.353*	PISA	0.328	PISA	0.284
O_C	-.373*	O_C	-0.271	O_C	-.434*	O_C	-.400*
N_C	-0.025	N_C	-0.025	N_C	0.163	N_C	0.165
MON	-0.082	MON	-0.07	MON	0.057	IP(eV)	-.450*
IP(eV)	-.593**	IP(eV)	-.549**	IP(eV)	-.462**	H_C	-0.197
H_C	-0.249	H_C	-0.262	H_C	-0.262	glob	0.088
FOSA	0.063	FOSA	0.094	FOSA	0.066	FOSA	0.131
FISA	0.129	FISA	0.214	FISA	0.25	FISA	0.273
E_L-E_H	-.868**	E_L-E_H	-.928**	E_L-E_H	-.913**	E_L-E_H	-.912**
EA(eV)	-0.248	EA(eV)	-0.105	EA(eV)	-0.103	EA(eV)	-0.125
donorHB	0.213	donorHB	0.191	donorHB	0.231	donorHB	0.243
dipole	0.124	dipole	0.135	dipole	0.15	dipole	0.157
DBE	.521**	DBE	.602**	DBE	.570**	DBE	.547**
C=C	0.314	C=C	.386*	C=C	0.346	C=C	0.307
Arom	0.228	Arom	0.29	Arom	0.262	Arom	0.232
accptHB	0.073	accptHB	0.088	accptHB	0.027	accptHB	0.05
#rtvFG	-.489**	#rtvFG	-.628**	#rtvFG	-.725**	#rtvFG	-.724**
#rotor	0.073	#rotor	0.1	#rotor	0.023	#rotor	0.056
#ringatoms	.467**	#ringatoms	.454**	#ringatoms	.418*	#ringatoms	.394*
#nonHatm	0.218	#nonHatm	0.266	#nonHatm	0.302	#nonHatm	0.317
#noncon	0.134	#noncon	0.069	#noncon	0.01	#noncon	0.027
#in56	.502**	#in56	.492**	#in56	.445*	#in56	.421*

**. Correlation is significant at the 0.01 level

*. Correlation is significant at the 0.05 level

B2: Correlations between molecular descriptors and rate constants of OMPs from *Chapter 3*

	lnk_{O3}		$k_{OH}(10^9)$
lnk_{O3}	1	$k_{OH}(10^9)$	1
X	-0.115	X	-.513**
WPSA	-0.096	WPSA	-.506**
volume	.385*	volume	.790**
SASA	0.379	SASA	.792**
QPpolrz	.415*	QPpolrz	.827**
PSA	-0.172	PSA	.737**
PISA	0.365	PISA	.712**
OH	0.021	OH	.525**
O_C	-.555**	O_C	0.19
N_C	0.242	N_C	.353**
MON	-0.262	MON	0.245
IP(eV)	-.866**	IP(eV)	-.704**
H_C	0.113	H_C	-.439**
FOSA	0.365	FOSA	0.124
FISA	-0.318	FISA	.705**
$E_L - E_H$	-.531**	E_L-E_H	-.424**
EA(eV)	0.202	EA(eV)	.463**
donorHB	0.159	donorHB	.620**
dipole	0.073	dipole	.520**
DBE	0.332	DBE	.950**
C=C	.456*	C=C	.789**
Arom	.423*	Arom	.670**
accptHB	0.216	accptHB	.675**
#rtvFG	-0.334	#rtvFG	-0.125
#rotor	0.255	#rotor	.653**
#ringatoms	.429*	#ringatoms	.876**
#nonHatm	0.345	#nonHatm	.855**
#metab	.386*	#metab	.728**
#acid	-.426*	#acid	0.003

**. Correlation is significant at the 0.01 level
*. Correlation is significant at the 0.05 level

APPENDIX C, Questionnaire

The following questionnaire was created and sent to the participants (academicians/industry) and the results were used to assign weights to the effects.

1. Which of the three factors are most influential during water treatment process? Rank them accordingly.
   - micropollutant removal
   - natural organic matter hindrance
   - by-product formation

2. What technical aspects are important in a water treatment plant?
   - reliability and maintenance
   - professional skill required
   - potential for modification

3. Rank the three environmental impacts accordingly.
   - waste and wastewater discharge
   - CO_2 emission
   - Land Use

4. With respect to resource use, rank the following.
   - energy use (power consumption)
   - chemical use

5. Which of the sustainability criteria do you consider significant.
   - environmental impact
   - resource use
   - public acceptance

6. In a water treatment process, rank the level of importance of the following.
   - treatability
   - treatment cost
   - technical consideration
   - sustainability
   - residence time