Prediction of Behaviour Kinetics and Toxicity of Engineered Nanomaterials in Aqueous Environment Using Neural Networks

By

Mary Nelima Ondiaka

Dissertation presented for the Degree

of

DOCTOR OF PHILOSOPHY

(Chemical Engineering)

in the Faculty of Engineering

at Stellenbosch University

Supervisor

Prof. Ndeke Musee

Co-Supervisors

Prof. Chris Aldrich Dr. Annie Chimphango

December 2016

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2016

Summary

As the global market for engineered nanomaterials (ENMs) continues to grow, the release of ENMs into the environment expose fauna and flora to new diverse stressors. Consequently, a novel approach is required in the monitoring and mitigation of the pollution from ENMs.

In this approach, modelling is likely to play a significant role in estimating the release, bioavailability, and toxicity of ENMs in the environment. Many laboratory tests have established the toxic effects associated with the acute and chronic exposure of various organisms to ENMs. However, most of the information generated from these tests and reported in the scientific literature is unstructured and uncertain. A further complication is that unlike with other pollutants, extrapolating experimental findings to assess the environmental risk of ENMs is difficult, owing to their diversity, lack of standardized test protocols and unknown allowable environmental concentrations. Dissimilar ENMs would require case-by-case risk evaluation, a process that would be expensive and time-consuming.

Focusing on $nTiO_2$ as a model ENM, and algae and *Daphnia magna* as indicator organisms, this dissertation presents learning from a database derived from information gathered from the scientific literature. More specifically, an ensemble model trained using multilayer perceptron neural network (MLP-NN) predicts mass coverage of organic adsorbates on $nTiO_2$, as well as the hydrodynamic size of $nTiO_2$ particles. These two response variables represent the behaviour kinetics of the particles in aquatic conditions. Also, the toxicity of the particles was predicted from selected characteristics of $nTiO_2$ and assay water, as well as some biological factors.

The neural network models could subsequently be interrogated to establish the effect of the various predictors of the behaviour of the particles associated with their environmental risk. This approach lays a foundation for the data mining of ENMs that would facilitate the use of available data to estimate the ecological impact of $nTiO_2$, as well as other ENMs.

Opsomming

Soos wat die globale mark vir ingenieursnanomateriale (INMe) aanhou om te groei, stel die vrystelling van INMe in die omgewing fauna en flora bloot aan nuwe diverse stressors. Gevolglik benodig die monitering en bekamping van besoedeling deur INMe ‘n nuwe benadering.

In die benadering, is dit waarskynlik dat modellering ‘n beduidende rol sal speel in die beraming van die vrystelling, biobeskikbaarheid en toksisiteit van INMe in die omgewing. Laboratoriumtoetse het die toksiese effekte bevestig wat gepaard gaan met die akute en chroniese blootstelling van organismes aan INMe. Meeste van die inligting wat deur die toetse gegenereer word en in die wetenskaplike literatuur gerapporteer word, is egter ongestruktureerd en onseker. ‘n Verdere komplikasie is dat anders as met ander besoedelingstowwe, is dit moeilik om vanaf eksperimentele bevindings te ekstrapoleer om die omgewingsrisiko van INMe te beraam, a.g.v. hulle diversiteit, gebrek aan gestandaardiseerde protokolle en onbekende toelaatbare omgewingskonsentrasies. *Daphnia magna* as indicator organisms

Deur te fokus op $nTiO_2$ as ‘n model-INM, en alge en *Daphnia magna* as indikatororganismes, bied hierdie proefskrif leer aan vanaf ‘n databasis wat afgelei is van inligting wat uit die wetenskaplike literatuur versamel is. Meer spesifiek, ‘n ensemble van veellaag-perseptron- neurale netwerke is gebruik om die akkumulasie van organiese adsorbate op $nTiO_2$, asook die hidrodinamiese grootte van $nTiO_2$ deeltjies te voorspel. Die twee responsveranderlikes is gebruik om die kinetiese gedrag van $nTiO_2$ te benader. Daarby is die toksisiteit van die deeltjies ook voorspel vanaf geselekteerde eienskappe van $nTiO_2$ en toetswater, sowel as ‘n aantal biologiese faktore.

Die neurale netwerkmodelle kon gevolglik geïnterrogeer word om die effek van die onderskeie voorspellers op die gedrag van die deeltjies wat met omgewingsrisiko geassosieer word, vas te stel. Die benadering lê die fondament vir die bevordering van data-ontginning van INMe wat die gebruik van beskikbare data om die omgewingsimpak van $nTiO_2$, sowel as ander INMe te beraam.

Acknowledgements

An endeavor to navigate uncharted waters requires good sail. For this to happen, I thank my promoters Prof. Ndeke Musee, Dr. Annie Chimphango and Prof. Chris Aldrich for their confidence in me, and incessant support and guidance to venture into an emerging and challenging field of study.

I thank Dr. Godfrey Madzivire for introducing me to hydro-geochemistry and guiding on PhreeQC modelling.

I am grateful to the Water Research Commission (WRC), Council for Scientific and Industrial Research (CSIR) and Stellenbosch University for providing me the financial support to pursue this study.

I am indebted to the staff and colleagues in the Department of Process Engineering at Stellenbosch University for providing a favourable environment that allowed me to focus on the study, and my friends for their moral support, good wishes, and encouragement. You permitted me to rejoice with you during good weather and lean on you when I hit the icebergs. Exceptional mention goes to all members of the international bible study group guided by Ulli and Heide Lehmann, Nathan and Jane Chiroma's family, Aline Uwimbabazi, Stephen Haingura, and Nusrat Begum for the priceless friendships and being part of my home away from home.

My gratitude goes to my entire family members for their unquestionable love, understanding, inspiration, and support.

A special dedication goes to my dear parents; the late Joash Ondiaka Marisio (2005) and the late Sarah Masitsa Ondiaka (1994). Your candles burnt out but the light you ignited still burns.

Lastly, I thank God for the gift of life.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

DLVO	Derjaguin-Landau-Verwey-Overbeek model
AlPDMSGI	$Al(OH)_3$, polydimethylsiloxane, glycerin and other organics capping agent
Al	Aluminium oxide capping agent
Algal TM	Standard algal test media
AlPDMS	Aluminium hydroxide and polydimethylsiloxane capping agent
AlSi	Aluminium and silicon oxides capping agent
BFGS	Broyden-Fletcher-Goldfarb-Shanno algorithm
CCC	critical coagulation concentration
CG	conjugate gradient algorithm
CLP	Classification, Labelling, and Packaging
CSIR	Council for Scientific and Industrial Research
DLS	dynamic light scattering
DOC	dissolved organic carbon
ECB	European Chemicals Bureau
ECHA	European Chemicals Agency
ECOTOX	ecotoxicology
EDL	electrical double layer
ENMs	engineered nanomaterials
EPM	electrophoretic mobility
EU	European Union
FA	fulvic acid
FIFFF	flow-field flow fractionation
GD	gradient descent algorithm
GW	ground water
HA	humic acid
HEPES	4-2(-hydroxyethyl)-1-piperazineethanesulfonic buffer
HW	synthetic hard water
IS	ionic strength
IsoLeF	Iso Lehmälampi freshwater
J24	< 24-hours juveniles *D. magna*
KDD	knowledge discovery from data
LaBW	Långskär brackish water

List of Abbreviations (Continued)

LM	Levenberg-Marquardt algorithm
LQ	lower quartile
LW	lakewater
MES	2-(N-morpholino) ethanesulfonic organic buffer
MHW	synthetic moderately hard water
MLP-NN	multilayer perceptron neural networks
MSE	mean square error
MWCNTs	multi-walled carbon nanotubes
N24	< 24-hours neonates *D. magna*
NOM	natural organic matter
NTU	nephelometric turbidity units
OECD	Organization for Economic Co-operation and Development
PBW	peat bog water
PDMS	polydimethylsiloxane
PZC	the point of zero charge
PZC	point of zero charge
QSAR	quantitative structure activity relationship
QSAR	quantitative structure-activity relationship
RDCM	reconstituted *Daphnia magna* culture media
REACH	Regulation, Evaluation, Authorization, and Restriction of Chemical
RQ	risk quotient
RW	river water
SA	sensitivity analysis
SAC	surface area concentration
SB_SW	Santa Barbara seawater
SCENIHR	Scientific Committee on Emerging and Newly Identified Health Risks
SFA	substance flow analysis
SGD	stochastic gradient descent algorithm
SimF	Simijärvi freshwater
SOM	surrogate organic matter
SRFA	Suwanee River fulvic acid
SRHA	Suwanee River humic acid

List of Abbreviations (Continued)

SRNOM	Suwanee River natural organic matter
SSA	specific surface area
SW	seawater
SWCNTs	single-walled carbon nanotubes
TC	total carbon
TOC	total organic carbon
TvBW	Tvärminne brackish water
TW	tap water
UQ	upper quartile
US EPA	United States Environmental Protection Agency
VHW	synthetic very hard water
VSW	synthetic very soft water
WRC	Water Research Commission
WWTPs	wastewater treatment plants

Chapter 1 - Introduction

1.1 Engineered Nanomaterials

Engineered nanomaterials (ENMs) are zero-, one-, two-, and three-dimensional nanostructures that measure ≤ 100 nm at least in one direction. Nanostructures include, but are not limited to, nanoparticles, nanorods, nanotubes, quantum dots, nanofilms, and nanocomposites. Novelty in nanostructures is their specialized design and engineered properties for specific application in various industries, such as healthcare, sports, food, and cosmetic among others.

The design and production methods determine the properties and quality of ENMs. The nanoscale size and surface imperfections of ENMs produced enhance their surface energies and reactivity. Moreover, process impurities may result in uncertain and dissimilar properties of ENMs. The quality of ENMs produced raises concerns in industry and environment as raw ingredients and pollutants requiring monitoring, respectively. Coating the surfaces of ENMs with organic and inorganic agents improves their quality and stability for specific applications. However, the surface coating would determine the environmental persistence and bioavailability of ENMs as altered pollutants. Therefore, positive aspects of ENMs desired by industry may be negative attributes after their release into the environment.

Physicochemical properties vary with dissimilar types and nanostructures of ENMs. Notable ENMs have metal, carbonaceous, and metal oxide origin, such as nano-silver (nAg), fullerenes, nano-cerium oxide (nCeO), nano-zinc oxide (nZnO) and nano-titanium dioxide ($nTiO_2$). Nano-silver is a renowned biocide (Nowack et al. 2011). The CNT are a group of single- and multiple-walled rigid fullerenes exhibiting high tensile and elasticity strength among thermal, electrical and other properties (Cadek et al. 2002; Coleman et al. 2006). Examples of CNT application include the manufacture of biomedical devices and energy storage systems that require materials having high strength and thermal conductivity, respectively.

The nCeO is a rare earth metal oxide with high catalytic properties relevant in enhancing combustion in the automotive industry (Nolan et al. 2006). There is a wide application of photocatalytic $nTiO_2$ and nZnO as pigment ultraviolet blockers in sunscreens and other cosmetic products. Bulk forms of TiO_2 and ZnO are used in photo-remediation of pathogenic and organic contaminants in water (Linsebigler et al. 1995; Fujishima et al. 2000; Hariharan

2006). The bulk TiO_2 that consists ENMs ≤ 100 nm is applied to food as an anti-caking agent, colourant, texture enhancer or surface shell owing to its opacity (Weir et al., 2012).

During their life cycle (production, manufacture, use, and disposal), the ENMs are released as environmental pollutants. In 2010, projected global production of the notable ENMs placed $nTiO_2$ ahead of nZnO, nCeO, CNT, and nAg, in the given order (Keller et al. 2013). The percent release into the surface water during use and disposal stages was $nTiO_2$ (53.41), nZnO (12.67), nCeO (1.03), nAg (0.22) and CNT (0.11) (Keller et al. 2013). The estimates of released ENMs were contingent on presumed application and matrix forms or location in commercial nanoproducts. For instance, the ENMs found in liquid suspensions (aerosols, cosmetics, and others) and bulk forms (nanopowders and others) would be released into air, soil, water, and wastewater collection systems. Besides, ENMs bound on surfaces, for instance, thin films, or embedded in products, such as food packaging materials, can be disposed of as solid wastes in landfills.

The list, descriptions, and applications of the ENMs described above are not exhaustive. Still, their application demonstrates expected disposal of CNT as solid waste in landfills, and release of nAg, nCeO, $nTiO_2$, and nZnO into the aquatic environment. Therefore, excluding CNT, the four ENMs qualify for investigation as case study aquatic pollutants, led by $nTiO_2$.

1.2 Relevance of Aquatic Environment as a Receptor of ENMs

The aquatic environment is a prospective reservoir for ENMs discharged directly or indirectly from air, soil, transient water, and wastewater treatment plants (WWTPs). Atmospheric depositions, terrestrial erosions, water runoff, and effluent contribute to the build-up of ENMs in surface water. The receiving aquatic ecosystems are complex considering the source, composition, properties, and associations of physical, chemical, and biological ligands present. After their discharge, an interaction of the aquatic ligands with ENMs would influence the latter's behaviour kinetics, bioavailability, and toxicity (Navarro et al. 2008).

The kinetic mechanisms of ENMs and functional networks of diverse aquatic ecosystems are fundamental processes that would concurrently influence exposure and bioavailability of pollutants, and uptake and effects to individual organisms and their communities. Once released into the aquatic environment, weathering, dissociation, speciation, and adsorption kinetics would modify properties of ENMs while aggregation, deposition, settlement, and transport kinetics would influence their fate and bioavailability. Thus, the buildup of ENMs in each timeframe and compromised water quality are prospective impacts after their release.

Moreover, the potential transformation of ENMs and creation of new pollutants may require new management strategies (Nowack et al. 2012), which may have financial implications, such as water treatment costs.

Diversity and dynamism of aquatic ecosystems facilitate self-regulation and conservation of core functions, for instance, recycling nutrients and sustaining food webs. Existence and survival of aquatic fauna and flora in their natural habitats depend on interrelated functional networks between biotic and abiotic factors. Biotic factors (such as organism species and predation) represent heterogeneous living components of aquatic ecosystems that directly or indirectly affect fauna and flora. Intra- and interaction of biological organisms, competition for resources and external human factors, such as the release of pollutants, affects biotic stability in ecosystems. The abiotic factors, such as the pH, temperature, radiation, salinity, and water flows, organisms' survival and influence the interaction and toxicity of pollutants.

The effects of chemical pollutants manifest at the cell, organism, population, community and ecosystem levels (Rand et al. 2003), a continuous hierarchy from the smallest to the highest. Besides, aquatic trophic levels (primary producers and consumers, and secondary and tertiary consumers) determine successive transfer of pollutants in the food chains and webs. Hence, bioavailability and effects of ENMs to cells would be sequentially expressed in the hierarchy and trophic levels at different stages of exposure leading to ecosystem impacts.

Though aquatic organisms physiologically adapt to certain levels of pollutants after exposure (Rand et al. 2003), pollution from new substances with unknown discharge concentrations like ENMs may disturb the core functions of ecosystems. For their dependence on tertiary consumers as a source of food, such as fish, humans may occupy the apex of an upright energy pyramid, a position demonstrating exposure to unknown risks of ENMs' pollution.

1.3 Nanoecotoxicology

In engineered systems, for example, water treatment plants, germicidal ENMs enhance the quality of water by killing pathogens. Potential natural treatment of pathogens by ENMs released in the aquatic environment is unknown. Although this may give the impression of positive attributes, the toxicity of ENMs to aquatic organisms and unknown long-term effects to ecosystem functions requires assessment and monitoring. Nanoecotoxicology is a developing field of study driven by the need to assess environmental fate, behaviour, and toxicity of ENMs. Environmental monitoring, and exposure, toxicity and modelling studies, are strategies adopted to address nanoecotoxicology issues of ENMs.

Analytical methods applicable for evaluating organic and inorganic compounds in the environment have been adopted for similar aspects in nanoecotoxicology (Abbott and Mightnard, 2010; Von der Kammer et al., 2012). Gottschalk et al. (2013) reviewed various techniques applied to determine the concentration of ENMs found in environmental samples. For instance, water runoff from exterior facades contained 0.001-0.175 mg/ℓ of nAg (Kaegi et al. 2010) and *ca.* 0.6-1.0 mg/ℓ of $nTiO_2$ (Kaegi et al. 2008). Moreover, the concentration of $nTiO_2$ determined in river water was found to be 0.01-0.011 mg/ℓ (Neal et al., 2011) and in effluent released into surface water was 0.01-0.015 mg/ℓ (Kiser et al. 2009), 0.042 mg/ℓ (Westerhoff et al. 2011) and 2.7×10^{-6} mg/ℓ (Khosravi et al. 2012).

Continuous collection of wastewater, treatment, and discharge of effluent into surface water makes WWTPs important conduits of ENMs. Analytical imaging methods reveal preserved intrinsic properties of ENMs in effluent (Kiser et al. 2009; Westerhoff et al. 2011). Concerns have been raised about likely compromised efficiency of WWTPs from effects of biocidal ENMs (Brar et al. 2010; Musee et al. 2011). The potential toxicity effects of ENMs on bacteria may reduce biodegradation capabilities (Musee et al. 2011) and enhance their release in effluent (Kaegi et al. 2011), hence, their accumulation in surface water.

Exposure studies link the adsorption and aggregation behaviour of ENMs to their physicochemical properties and simulated abiotic factors in natural water (Keller et al. 2010; Ottofuelling et al. 2011; Chekli et al. 2015; Chekli et al. 2015). Similar findings were established using standard synthetic or purified aqueous media (Domingos et al. 2009; Hartmann et al. 2010; Thio et al. 2011; Campos et al. 2013). The behaviour kinetics of ENMs influence their interaction with biological organisms in an aqueous environment (Battin et al. 2009; Horst et al. 2010). Conversely, established filtration of ENMs in porous media elucidates hindered transport in subsurface water (Chowdhury et al. 2011), which suggest negligible releases of ENMs into the surface water via groundwater gravity flows.

The first dose-response study by Oberdörster (2004) revealed antibacterial effects and toxic degradation of lipids in fish after acute exposure to 0.5 mg/ℓ of Fullerenes (nC_{60}). Since then, similar studies have used various testing protocols to demonstrate acute and chronic ecotoxicity of a wide range of ENMs to diverse aquatic organisms, such as algae, bacteria, daphnia sp., fish, worms, snails, yeasts, aquatic midges, shrimps, and higher plants among others. Findings presented in reviews (Kahru and Dubourguier 2010; Dhawan and Sharma 2010; Hou et al. 2013; Von Moos et al. 2014) and perspectives (Boxall et al. 2007) demonstrate high ecotoxicity effects of ENMs not comparable to their bulk forms. Under

diverse laboratory controlled systems, properties of ENMs contributed to observed toxicity effects. The uptake and bioaccumulation of ENMs in organisms underscored the potential long-term effects on their offspring and ecosystems.

Reviews (Gottschalk et al. 2013; Hendren et al. 2013) outline deterministic and stochastic modelling approaches applied or yet applied in nanoecotoxicology. Large-scale substance flow analysis and multimedia fate models estimate concentration, transport, fate and risks of ENMs in environmental compartments (Boxall et al. 2007; Mueller and Nowack 2008; Gottschalk et al. 2009; Musee 2011b; Blaser et al. 2007; Praetorius et al. 2012; Liu and Cohen 2014). Conversely, small-scale computational models estimate cause and effect relationships of ENMs and endpoints (Toropov and Leszczynski 2006; Barnard 2009; Puzyn et al. 2011; Apul et al. 2013; Liu et al. 2013; Pathakoti et al. 2014; Liu et al. 2014; Goldberg et al. 2015). The sturdiness and replicability of the models are contingent on the quality of available data; geospatial-temporal sites; assumptions used; and selected input variables.

This study applies multilayer perceptron (MLP) neural networks model to learn from collated scientific domain data on behaviour kinetics and toxicity of ENMs and train predictive models. The data used were generated using natural, standard synthetic and purified water under controlled conditions simulating the aquatic environment. The MLP is a feed-forward backward-propagation computational machine learning technique that interpolates and transforms information in databases to estimate nonlinear associations between predictors (or input variables) and target responses. Stochastic nature of this technique permits favourable conversion of uncertain data to learn important predictors, establish significant relationships, and provide infinite solutions aggregated into suitable finite prediction models.

The flexibility of neural networks favours potential application as a management tool in research and development to predict behaviour kinetics and toxicity effects of diverse types and nanostructures of ENMs. Conceptually, transformed properties of dissimilar ENMs can be either model inputs or responses in simulations and predictions that explain phenomena in the aquatic environment. The tool is adaptable to changes in external information by adjusting weights, which define the strength of inputs, in forecasting outputs of new related phenomena.

1.4 Problem Statement

Environmental risks associated with ENMs are necessary to evaluate but difficult to conduct because of limited understanding of their occurrence, exposure, and impacts. Information

available on global production volumes of ENMs and their industrial application exhibit discrepancies. Actual data on ENMs produced and used in nanoproducts are not documented or monitored owing to confidential business proprietary issues (The Royal Society and the Royal Academy of Engineering 2004). Moreover, it is not mandatory to disclose identities, properties, and content of ENMs in nanoproducts. Production volumes are estimated using available data and assumptions that can be subjective. Thus, informed decision-making by relevant stakeholders on the use of nanoproducts and disposal of nanowaste is hindered by the lack of substantive information on production volumes, types, properties and quantities of component ENMs (Musee 2011a).

The multistage life cycle of ENMs increases their potential risk profiles to different environmental matrices (Nowack and Bucheli, 2007). Positive application of enhanced ENMs contrasts their effects when directly and indirectly released into the aquatic environment as new pollutants during their life cycle stages (Fig. 1.1).

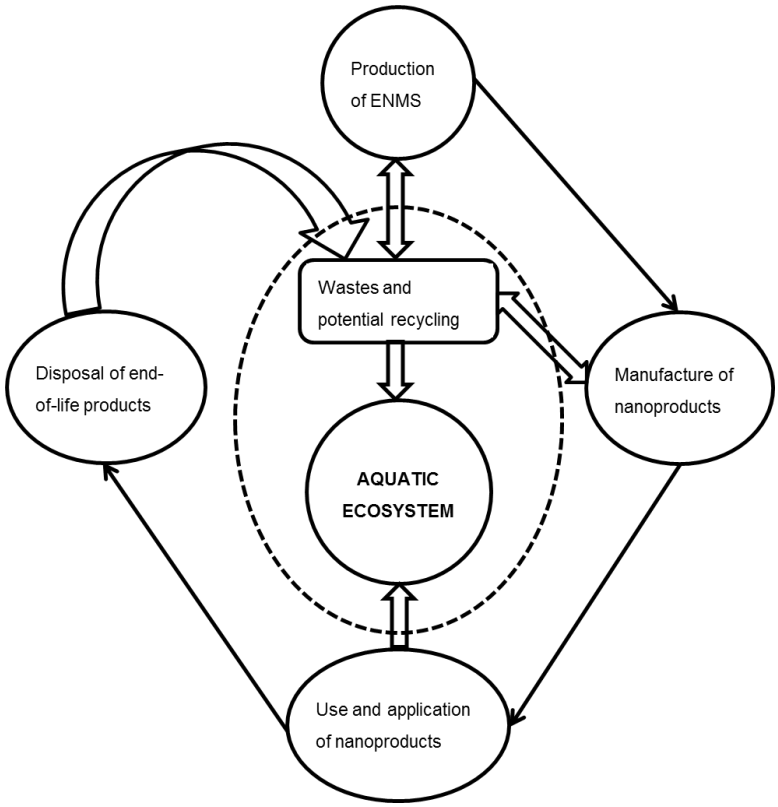


Figure 1.1 Illustration of potential life cycle release of ENMs into the aquatic environment.

There is inadequate information on short-, medium- and long-term environmental implications of ENMs to ecosystems, and eventually humans. Inadequate understanding of unintended consequences of nanoproducts surpasses the ability to assess their potential

release, exposure, and effects. Risks from nanotechnology are largely unknown to manufacturers, consumers, waste regulators, scientists, government officials, and other stakeholders at various stages of the value chain (Helland et al., 2008; Abbott and Mightnard, 2010).

There is no specialized in-situ equipment for analyzing ENMs in the environment. Besides, stoichiometry methods limit quantitative analyses of elemental concentrations of ENMs. However, analytical imaging techniques can distinguish natural and engineered nanomaterials in environmental samples. The lack of standard testing protocols, and diverse types and nanostructures of ENMs challenge the certainty and efficacy of experimental approaches and inference of findings as a basis for regulation. Authors have suggested an improvement to the current experimental methods in use (Handy et al. 2012).

Large-scale models in nanoecotoxicology rely on market information to estimate inputs and outputs of ENMs in the environment. Insufficient value chain information of ENMs, inconsistencies in estimated production estimates and lack of mandatory legislation on labelling (Gruère, 2011) for most products enhances model uncertainties. Simplified assumptions of analogous properties of ENMs and environmental systems may under- or over-estimate concentrations and potential impacts. Current application of descriptors that are more suitable for bulk materials hinders versatility of small-scale QSAR models. Authors (Puzyn et al. 2009; Xia et al., 2010; Hristovski et al. 2011; Westerhoff and Nowack 2013) suggest the development of nano-descriptors for meaningful nano-QSARs modelling.

The transformation, interaction with organisms, effects, and attenuation of ENMs in the aquatic environment is currently unknown. To estimate environmental risks of ENMs, one requires a good knowledge about actual or estimated concentrations and certain or estimated threshold levels. The inconsistency of reported risks for similar ENMs originated from diverse estimations of environmental concentrations and uncertain ecotoxicology data used (Mueller and Nowack 2008; Gottschalk et al. 2010; Musee 2011b; Nota 2011; Money et al. 2012; Gottschalk et al. 2013). Thus, there is no sound foundation for quantifying emissions, exposure and biological effects of ENMs applicable to aquatic ecosystems.

1.5 Motivation of Study

Technology-push innovations are industry-driven, for example, nanotechnology, which applies science and engineering principles to produce and use ENMs in various products. Consumers provide a negligible contribution to technical designs of ENMs, but form societies

carried by and depend on resources of, the environment. Unknown environmental risks of primary and transformed forms of ENMs pose mitigation and management challenges. Assessing behaviour kinetics and toxicity of ENMs informs about the likely human and environmental effects. The following factors delimit the release of nanowaste into the environment

- Techniques applied to produce ENMs as raw materials for nanoproducts.
- Methods used to pretreat raw materials and manufacture nanoproducts.
- Environmental, health, and safety protocols adopted by institutions and industries.
- Regulations and legislations on hazardous chemicals and wastes.

Characterization, management of nanowaste streams and moderation of environmental risks of ENMs are important at the production and manufacturing levels. The scale of production, synthesis methods, pretreatment techniques and quality of ENMs produced govern quantities and properties of nanowaste generated. Poor quality ENMs could be recycled or discarded (illustration in Fig. 1.1). The release of ENMs applied in nanoproducts may result in involuntary risks in the aquatic environment.

Standard operating procedures and work practices implemented in research laboratories and processing plants minimize unintentional exposure and prevent environmental hazards of ENMs. However, accidental exposures to ENMs during use and disposal are likely. For instance, down the drain collection of ENMs released from personal care products, cosmetics, and paints into freshwater. Limited studies have reported potential environmental weathering of nanoproducts and release of ENMs (Botta et al. 2011; Virkutyte et al. 2012). Accumulation of unquantifiable amounts of ENMs poses human and environmental risks.

Hazard, dose-response and exposure assessments, and risk characterization constitute procedural aspects of environmental risk assessment (Mihelcic and Zimmerman 2010; Ted 2014), which are adopted in nanoecotoxicology and described in Section 1.3. The evaluation of hazardous substances safeguards the well-being of exposed flora and fauna in ecosystems, and in turn humans. Thus, evidence of ENMs' risks from scientific research would be a basis for setting regulatory standards aimed at protecting humans and the environment. A good knowledge about properties, exposure, bioavailability, and ecotoxicity of ENMs is required to support effective regulation. In the absence of standard testing protocols for ENMs, risk assessment and regulation is challenged by the following,

- Dissimilar ENMs, nanostructures, and nanoproducts nullify 'one size fits all' regulation.

- Absent definite quantifiable or unquantifiable indicators of risks to support control measures.
- The indistinct difference between innovativeness and enhancement in nanotechnology for nanomaterials that have been in use for a long time.

Regulation enables business and environmental monitoring of ENMs along the value chain of nanotechnology, which would inform consumers about benefits and alert them on potential hazards. Besides the European Union (EU) that has instituted regulatory measures for ENMs, regulation agencies classify ENMs as hazardous chemicals in existing laws (Aschberger et al. 2014). The broad regulation approach does not provide guidelines for assessing ecotoxicity of ENMs. Under Regulation, Evaluation, Authorization and Restriction of Chemical (REACH) and Classification, Labelling and Packaging (CLP), regulation of ENMs by European Chemicals Agency (ECHA) emphasizes registration of products, categorization and labelling of ENMs among others (ECHA 2015). Regulation under REACH aims at controlling hazards, creating awareness, and protecting consumers of nanotechnology products in EU. Specific provisions for regulating ENMs ≤ 100 nm have been approved for biocide products as well as their risk management (ECHA 2015).

A strict proposal by United States Environmental Protection Agency (US EPA) to control production, processing, and importation of ENMs was replaced by a soft stance version that recommends monitoring and documentation (EPA 2015). Australia and New Zealand adopted EU's regulations but proposed modifications to incorporate regulation of ENMs in food (Fletcher and Bartholomaeus 2011). The absence of distinct variations between novelty and enhanced nanoproducts had led to exemptions of $nTiO_2$ and nZnO from regulatory labelling and monitoring in food and cosmetic products. However, the proposals by New Zealand's Environmental Protection Authority required mandatory labelling of $nTiO_2$ and nZnO in cosmetics (Moore 2012) that would also regulate their use and commercialization. In developing countries, there is no report outlining measures adopted to regulate ENMs.

Labelling and monitoring as a regulatory standard for ENMs have demerits. Studies show that nanoproducts are voluntarily labelled to disclose the content of ENMs (Gruère, 2011). For example, in 2014, inventories reported holding 49.0% (Project on Emerging Nanotechnologies 2015) and 66.0% (Danish Consumer Council 2015) of unlabeled nanoproducts that claim to contain ENMs. A study revealed a two-third probability of labelling the contents of ENMs in nanoproducts (Lomer et al., 2000). Another study established 36.0% of nanoparticles ≤ 100 nm in bulk ENMs preferentially applied in food and

cosmetics without adhering to concentration standards (Weir et al., 2012). The disparities between labelled and analyzed contents of ENMs in nanoproducts reduce accountability in the production value chain of nanotechnology, and capacity to monitor associated risks.

Nanowaste forms part of waste streams into existing disposal and treatment systems. Legislation on environmental wastes, implementation, and enforcement vary with regions. Reducing or preventing industrial nanowaste requires re-designing materials and processes, recovery and recycling strategies, among others. Since information on production is largely unavailable, these measures are not defensible at this stage. Treatment efficiencies of WWTPs determine removal of ENMs, especially as adsorbed solids, and release in effluent. Wastewater analyses reveal sludge adsorption interaction that removes 95.0-99.5% of $nTiO_2$ (Kiser et al. 2009). However, competency of WWTPs does not demonstrate mass flow rates and build-up of ENMs released into the aquatic environment.

Disposal of sludge on landfills or application in agriculture transfers unknown quantities and risks of ENMs in soils. Impacts of thermal transformation of ENMs during incineration and disposing of fly ash are also unknown. In countries with poor legislation on waste management, consequences of nanowaste that by-pass treatment, end-of-life disposal of nanoproducts in landfills, and reuse (for example, food packages coated with ENMs) are unidentified. Likely movement of nanowaste from landfills to aquatic environment is unknown. Uncertainties in nanotechnology inhibit environmental legislation of nanowaste.

Investigating the environmental exposure, bioavailability, toxicity, and effects of ENMs at the initial stages of nanotechnology development is important. However, it is not prudent to investigate specific behaviour and toxicity for each type and nanostructure of ENMs. Hence, data-driven modelling is suitable for extracting knowledge from domain data to elucidate relationships between causal factors and behaviour kinetics and toxicity of diverse ENMs in an aquatic environment. Authors suggested characterization of ENMs and aqueous environment as a means to supporting predictive analytics in nanoecotoxicology and application of findings (Nowack and Bucheli 2007; Christian et al. 2008; Klaine et al. 2008; Hornyak et al. 2009; Stone et al. 2010; Lin and Tian 2010; Petosa et al. 2010; Musee et al. 2011; Batley et al. 2013). From a nanotechnology value chain and environment perspective, long-term benefits of predicting the behaviour and toxicity of ENMs includes the following,

- Saving time and costs for analyzing multiple ENMs, nanostructures, and responses.
- Supporting decisions made for safe production and application of nanomaterials.

- Supporting research, which would promote responsible nanotechnology for a healthy environment.

1.6 Research Questions

This research interrogates published scientific reports to establish inherent physicochemical properties of ENMs, and environmental abiotic and biotic factors that significantly influence their exposure and biological effects in the environment. In turn, these factors are applied to predict behaviour kinetics and toxicity of ENMs. Specific research questions include:

1. What inherent physicochemical properties of ENMs, environmental abiotic and biotic factors influence behaviour kinetics of ENMs in an aquatic environment?
2. What are the relationships between inherent physicochemical properties of ENMs and environmental abiotic and biotic factors, and behaviour kinetics and biological toxicity?
3. To what extent does behaviour kinetics influence bioavailability and consequent toxicity of ENMs to aquatic organisms?

1.7 Study Objectives and Scope

The main objective of this study is to train models that predict behaviour kinetics of ENMs and toxicity effects on aquatic indicator organisms. Specific objectives include the following,

1. Establish inherent physicochemical properties of ENMs, and environmental abiotic and biotic factors influencing behaviour kinetics and biological toxicity during exposure.
    - Critically examine published scientific reports to excerpt and collate information on behaviour kinetics and biological toxicity of ENMs in an aqueous environment.
    - Assemble a training dataset that combines predictor and response variables from different scientific studies on behaviour kinetics and toxicity of ENMs.
2. Investigate relationships between physicochemical properties of ENMs, and environmental abiotic and biotic factors, and behaviour kinetics and biological toxicity effect responses using multilayer perceptron neural networks (MLP-NN) model and ensemble prediction paradigm.
3. Evaluate the associations between behaviour kinetics and biological toxicity effect responses using the fitted neural networks models.

The scope of this study is limited to utilizing the MLP-NN model to learn from labeled stratified scientific data on behaviour kinetics and toxicity of ENMs. Selected physicochemical properties of ENMs and environmental abiotic and biotic factors will be model inputs. Testing and validation of trained solutions using previously unseen data will measure performance and accuracy of the fitted models in predicting new phenomena. Sensitivity analyses of finite ensembles combined using aptly trained models will establish important factors governing behaviour kinetics and toxicity of ENMs, approximate input-output relationships and simulate discrete continuums of phenomena to test applicability.

This study will focus on investigating adsorption and aggregation behaviour of $nTiO_2$ and in turn, toxicity effects on algae and water flea, *Daphnia magna*, as model ENM and biological organisms, respectively. Projected high production and release of $nTiO_2$ into surface water requires an understanding of likely liquid-solid phase interactions. These linkages define partitioning with aquatic ligands, the alteration of surface chemistry, fate, and bioavailability of ENMs. Algae and *D. magna* are major ecotoxicity indicator organisms that are primary producers and consumers, respectively, significant in trophic transfer of pollutants. Biological toxicity endpoints are limited to biomass growth in algae and reproduction in *D. magna*.

1.8 Dissertation Layout

Chapter 2 presents an evaluation of nanotechnology and modelling approaches undertaken to address nanoecotoxicology aspects related to behaviour kinetics and toxicity of ENMs. Chapter 3 provides an overview of neural networks modelling and describes ensemble-modelling paradigm applied in this study. Discussed in Chapter 4 is a critical examination of published scientific literature on behaviour kinetics and biological toxicity of $nTiO_2$ and extraction of data, a compilation of a sample dataset for learning, and identification of model inputs and targets.

Chapters 5, 6 and 7 present modelling experiments, predictions, and interpretation of findings. Chapter 5 covers prediction of organic adsorbates on the surface of $nTiO_2$ under solid-liquid interaction regimes. Chapter 6 investigates prediction of the hydrodynamic size of $nTiO_2$ during exposure and effect of organic adsorbates on aggregate sizes of $nTiO_2$. Chapter 7 describes the prediction of toxicity effects of $nTiO_2$ on algae and *Daphnia magna* and influence of behaviour kinetics to toxicity. Chapter 8 revisits research questions and objectives to summarize findings, draw conclusions, and suggest areas requiring further studies. Attached in the Appendix is information relevant to the study.

Chapter 2 - Literature Review

Chapter 1 introduces challenges in nanoecotoxicology. This Chapter highlights nanotechnology issues concerning potential release and environmental risks of engineered nanomaterials (ENMs) and principles that govern behaviour kinetics of charged solids in an aqueous environment. The modelling approaches evaluated are those that focused on mass concentration aspects to address nanoecotoxicology of ENMs. Knowledge gained illuminates on relevant model inputs and implementation in this study.

2.1 Nanotechnology

Taniguchi (1974) coined the term 'nanotechnology' to describe the precision reduction of bulk materials to nanometre range in thin film passivation of semiconductors. The current use of the terminology describes various scientific and technological research, deliberate design, specialized manipulation, characterization and application of atoms and molecules at the nanometre scale (Cao 2004; Bowman and Hodge 2007; Casals et al. 2008). Specialized engineering techniques have yielded nanoscale materials, devices, products, and systems having unique structural configurations. With ENMs as basic building blocks, two- and three-dimensional molecular assemblages are controlled to form well-defined nanostructures (Rosi and Mirkin 2005).

Scientific concession on the size of ENMs is accepted as nanostructures that are 1-100 nm at least in one dimension (SCENIHR 2007; Lövestam et al. 2010; EU 2011; Roebben et al. 2014). Sizes up to 1,000 nm have been advocated to include clusters or polydisperse ENMs (Lövestam et al. 2010), which incorporates those commonly used in nanomedicine (Mitsiadis et al. 2012). Complex structures, properties and behaviour of ENMs produced are comparatively uncommon with their bulk forms (Cao 2004; Hansen et al. 2007; Crane et al. 2008; Navarro et al. 2008; Casals et al. 2008).

Nanomaterials are centuries old and incidentally, occur in the environment (Hornyak et al. 2009). Medieval artistry related to nanotechnology produced historical masterpieces using complex organic and inorganic materials. Notable examples of age-old nanomaterials and products include the following,

- Iron carbide particles used in the 3rd to 17th century Damascene steel blades (Reibold et al. 2009).

- Colloidal gold and silver found in the 4th century 'Lycurgus Cup' (Hornyak et al. 2009).
- Colloidal nAg (≤ 10 nm) applied in biocides for more than a century (Nowack et al. 2011).
- Pigment TiO_2 produced for industrial application since 1916 (Jovanović 2014), which is a bulk ingredient that contains ENMs ≤ 100 nm (Weir et al., 2012).
- Copper oxide nanoparticles used in the 6th to 15th centuries' glass windows (Kunicki-Goldfinger et al. 2014).

In 1959, Richard P. Feynman conceptualized engineering of matter to atomic and molecular levels (Feynman 1992). The notion of ENMs spurred research and development that may be referred to as 'modern nanotechnology' after the invention of advanced equipment in the 1980s, for example, the scanning tunneling microscope in 1981 (Binnig and Rohrer 1987). Since the beginning of the 1990s, nanotechnology has matured from laboratory-based research and development phase into full commercialization of nanoproducts. At the start of 2000, advancement from elementary-to-complex nanotechnology has seen innovative products being commercialized globally (Roco et al., 2010). The novelty of nanotechnology is the production of anthropogenic ENMs with atypical characteristics that do not have equivalent natural forms (Cao 2004). The advancement of nanotechnology innovations raises concerns of ENMs as new pollutants introduced into the environment.

2.1.1 Production of ENMs

Production methods, which include synthesis, purification, functionalization and characterization processes, affect intrinsic physicochemical properties of ENMs found to be suitable for nanotechnology-derived products. Synthesis includes top-down (subtractive), bottom-up (additive) and hybrid methods that influence sizes, shapes (for example, spherical, rod- and needle-like, rings, tubes and films) and quality of ENMs (Cao 2004; Hornyak et al. 2009; Meyer et al. 2009). Top-down techniques, for instance, attrition and milling reduce bulk materials to the nanoscale level under controlled conditions. Bottom-up techniques, for example, sol and sol-gel assemble ENMs from precursor materials by controlling growth, homogeneity, and surface properties.

Sizes, structural defects, irregularity, and heterogeneity of ENMs produced by top-down methods can enhance surface energy and reactivity. Additive methods improve weaknesses of subtractive methods, although process contaminants affect the quality of ENMs. Poor quality ENMs yield undesirable raw materials that may have different toxicity effects when released into the environment from nanoproducts. Studies show that contaminants impact

surface characteristics (Liu et al. 2011), enhance conductivity (El Badawy et al. 2011) and induce observed biological toxicity (Hull et al. 2009). Purification improves the quality of ENMs by removing process contaminants, for example, washing in aqueous deionized water (Forough and Farhadi, 2010).

Functionalization creates distinct properties of ENMs for specific applications using physical, chemical, and biological grafting techniques. Surface passivation and doping in thin films (Fan and Lu 2005; Chen and Mao 2007; Xu et al. 2012), polymer/clay embedment (Duncan 2011) and capping (Wiley et al. 2007) are examples of functionalization that modify the surface structure and characteristics of ENMs. Strengthened nanostructures can influence disposal methods, release pathways, kinetic transport, fate, and persistence in the environment. Hence, knowledge about grafting materials used is important when monitoring ENMs released into the environment.

Characterization authenticates nanoscale properties of ENMs to evaluate quality for intended application. From an environmental perspective, ENMs' properties broaden the understanding of likely speciation, bonding, transformation, fate, bioavailability, and toxicity that facilitate nanoecotoxicology assessment. Table 2.1 illustrates examples of analytical techniques used to characterize various aspects of ENMs. Instruments used, accuracy and precision levels, material handling methods and detection limits affects characterized properties of ENMs (Weibel et al. 2005). The ENMs tend to exhibit different properties from the same batch (Ohtani et al. 2010; Akin et al. 2015), an attribute that may deter determination of definite attributes necessary for subsequent monitoring in the environment.

Structural orientation, nanostructure size and location, and risk profiles are categorization frameworks proposed for ENMs (Table 2.2). These frameworks overlap without independently defining ENMs. However, risk profile classification differentiates intrinsic and transformed properties of ENMs in the aquatic environment.

Table 2.1 Examples of methods used for characterizing and detecting ENMs

Examples of technique	Application	References
1. Optical probe and lesser scattering technique		
• Dynamic light scattering (DLS) • Small-angle neutron scattering (SANS) • Quasi-elastic light scattering (QELS)	• Particle and molecule size / distribution in liquids • Surface-to-volume ratio of crystals or amorphous particles • Particle size distribution	(Hornyak et al., 2009)
2.Electron probes technique		
• Scanning electron microscope (SEM), coupled with energy-dispersive X-ray (EDX) or energy dispersive spectroscopy (EDS) • Transmission electron microscope (TEM), coupled with EDX or EDS	• Surface morphology and microstructure, chemical composition (EDX and EDS), and distribution of nanomaterial, nanostructures, and bulk materials • Surface morphology, crystal structure and defects, the chemical composition of nanostructures and bulk nanomaterials; electrical and mechanical properties of nanostructures, and others.	(Hornyak et al. 2009; Cao 2004; Weibel et al. 2005)
3. Scanning probe technique		
• STM (Scanning tunneling microscopy) • AFM (Atomic force microscopy)	• 3D surface structure images, roughness, defects, size, and conformation of atoms and molecules. • 3D topographic mapping, the crystal structure of small particles, epitaxial and highly textured thin films; size, volume-mass and surface area distribution, and others.	(Hornyak et al. 2009; Cao 2004; Binnig and Rohrer 1987)

Table 2.1: Examples of methods used for characterizing and detecting engineered nanomaterials (*continued*)

Examples of technique	Application	References
4. Spectroscopic technique		
• X-ray diffraction (XRD)	• Crystal structure analysis of small particles, epitaxial and highly textured thin films; identification of unknown materials	(Cao 2004; Hornyak et al. 2009)
• Fourier transformation infrared spectroscopy (FTIR)	• Qualitative atomic arrangement, chemical bonds among others	
• Surface enhance Raman spectroscopy (SERS)	• Atomic structure, chemical identification, and quantification	
• X-ray fluorescence (XRF)	• Surface analysis (<~1.5 nm) and analysis of atomic structure.	
• X-ray photoelectron spectroscopy (XPS)	• Surface analysis, elemental composition	
• Inductively coupled plasma spectroscopy (ICP), coupled with atomic emission spectroscopy (AES), atomic absorption spectroscopy (AAS), optical emission spectroscopy (OES)	• Element analysis and chemical speciation of functional groups.	
• Ultraviolet-visible spectroscopy (UV-Vis)	• Chemical analysis	
• Small-angle x-ray scattering (SAXS)	• Size and geometry of particles/mesopores (1-100 nm)	
5.Ion particle/spectrometry probe methods		
• Mass spectrometry (MS)	• Qualitative and material structure analysis	(Cao 2004; Hornyak et al. 2009)
• Secondary ion mass spectrometry (SIMS)	• Surface analysis (<~1.5 nm) and depth profiling	
6.Thermodynamic methods		
• Thermal gravimetric analysis (TGA)	• Change in mass, purity, reaction rates, thermal stability	(Brunauer et al. 1938; Cao 2004; Weibel et al. 2005; Giesche 2006; Hornyak et al. 2009)
• Mercury porosimetry (nitrogen adsorption) (MP)	• Relative surface area, pore volume [mesopores (2-50nm), microspores (<2 nm)], pore density and size distribution	
• Brunauer-Emmet-Teller (BET)	• Relative surface area analysis	

Table 2.2 Examples of categorization frameworks for ENMs

Framework and Reference	Category 1	Category 2	Category 3	Category 4
Structural orientation (Cao 2004)	**Zero-dimensional:** Nanocrystals, poly-nanocrystals, amorphous particles, for example, spheres, cubes, quantum dots among others	**One-dimensional:** Nanowires, nanorods, and nanofibres.	**Two-dimensional:** Thin films	**Special nanomaterials (purely anthropogenic):** Carbon fullerenes and nanotubes, micro- and mesoporous structures, for example, zeolites, core-shell structures among others
Nanostructure (Casals et al. 2008)	**Nano-clusters:** Amorphous semi-crystalline with one dimension 1-10 nm	**Nano-powders:** Agglomerates of non-crystalline sub-units with one dimension < 100 nm	**Nano-crystals:** Single crystalline subunits with one dimension ≤ 100 nm, for example, metals, quantum dots, core-shell nanoparticles	
Location of nanostructure (Hansen et al. 2007)	**Nano-bulk structures:** Uniform materials, mixed materials (for example, nanoporous and copolymers)	**Nano-surface structures:** Structured on similar substrate, un-patterned film on different substrate, patterned film on similar substrate	**Nanoparticles:** Surface-bound, liquid suspensions, embedded in solids, air particulates	
Risk profiling (Nowack et al. 2012)	**Pristine ENMs:** raw forms of nanomaterials produced	**Product-modified ENMs:** functionalized nanomaterials for specific applications	**Product-weathered ENMs:** transformation of ENMs encapsulated in nanoproducts	**Environmentally transformed ENMs:** modified forms of the ENMs

Gradual increases in estimated quantities of ENMs produced annually (Table 2.3), and listed number of nanoproducts in public inventories (Project on Emerging Nanotechnologies 2015; Nanowerk 2015) indicates an expanding global market of nanotechnology. Despite inconsistent estimates of ENMs produced annually since 2002-to-date, reports show sustained high annual production of $nTiO_2$ from 2005 compared with other common ENMs (Table 2.3). Parallel estimates of nAg, nCeO, carbonaceous ENMs and nZnO produced between 2010-2011 were 0.5-1.8%, 1.8-11.4%, 0.02-10.0% and 18.0-39.0% that of $nTiO_2$, respectively (Keller et al. 2013; Piccinno et al. 2012)(Table 2.3).

Table 2.3 Estimated global production of ENMs (tons/year)

ENMs type\Year	'02	'03	'04	'05	'06	'07	'08	'09	'10	'11
nAg				4[5]			300-800[9] 434[5] 563[5]	5[5] 1,230[11]	452[6]	55[12]
nCeO							10,000[8]	10,000[8]	10,000[6] 10,000[8]	55[12]
CNT	0.003[1]	3[2]	100[2] 473[3]	295[7]		278[5]	426[3] 500[2]		3,200[6]	300[12]
Fullerene	0.15[1]			10[4]			5[5] 300[8]	300[8]	300[8]	0.6[12]
$nTiO_2$				2,000[10]	5,000[8]	5,000[8]	5,000[8]	3,000[5] 5,000[8]	5,000[8] 88,000[6]	3,000[12]
nZnO					20[8]	20[8] 528[5]	18[5] 1,800[5] 9,845[5]		34,000[6]	550[12]

[1]Plasmacarb 2003; [2]Borm et al. 2006; [3]The Royal Society and the Royal Academy of Engineering 2004; [4]Fortner et al. 2005; [5]Gottschalk et al. 2009; [6]Keller et al. 2013; [7]Köhler et al. 2008; [8]UNEP 2007; [9]Mueller and Nowack 2008; [10]US EPA 2010a; [11]Musee 2011b; [12]Piccinno et al. 2012.

2.1.2 Application and Release of ENMs

Public inventories (Project on Emerging Nanotechnologies 2015; Nanowerk 2015) document metal, metal oxides and carbonaceous nanostructures commonly used in nanoproducts. Presently, there are numerous consumer nanoproducts and industrial applications of

nanotechnology, for example, nano-electronics, tissue engineering, biomedicine, cosmetics, paints, pesticides, and water purification modules, among others. As an illustration, the number of catalogued nanoproducts increased from 54 in 2005 to 1,814 in 2014 (Vance et al. 2015), and from 1,212 in 2012 to 2,231 in 2015 (Danish Consumer Council 2015). Table 2.4 demonstrates common ENMs in nanotechnology compared and contrasted in subsequent texts for illustration purposes.

Table 2.4 Demonstration of ENMs commonly used in nanoproducts

ENM	Application examples	Reference
nAg	Biocide agent in nanomedicine, textiles, water and food.	(Chen and Schluesener 2008; Schmid and Riediker 2008; Hauri and Niece 2011)
nCeO	Diesel fuel additive.	(Cassee et al., 2011)
CNT	Nanomedicine, polymers, electronics and energy storage.	(Baughman et al. 2002; Bianco et al. 2005)
Fullerene	Waste treatment, nanomedicine, electronics and energy storage.	(Plasmacarb 2003; Hendren et al. 2011)
$nTiO_2$	Cosmetics, personal care products, paints, varnishes, coatings, food additive, electronics, and wastewater treatment.	(Lomer et al. 2000; Borm et al. 2006; UNEP 2007; Schmid and Riediker 2008; Robichaud et al. 2009; US EPA 2010a)
nZnO	Cosmetics, sunscreens, biocide, ceramics, rubber processing, wastewater treatment, and electronics.	(Schmid and Riediker 2008; Schilling et al. 2010; Rajendran et al. 2010; Wong et al. 2010)

In 2010, comparative global production of nAg, CNT, nCeO and nZnO was 0.5%, 3.6%, 11.4% and 39.8% that of $nTiO_2$ (Keller et al. 2013). Contrary to production volumes, metal nAg is a popular component in nanoproducts followed by $nTiO_2$, carbonaceous ENMs, and nZnO, in that order (Project on Emerging Nanotechnologies 2015). For example, in 2013, listed nanoproducts contained 2- and 11-fold nAg compared to $nTiO_2$ and nZnO, respectively (Project on Emerging Nanotechnologies 2015), corroborated by a 3- and 15-fold related listing (Danish Consumer Council 2015). The popularity of nAg may be attributed to regular cataloguing as a renowned biocide, and regulations requiring labelling and documenting biocidal nanoproducts (Vance et al. 2015).

From an estimated volume of ENMs produced, 0.4-7.0% was released into surface water depending on the application, matrix form and release pathway (Keller et al. 2013). Based on information from 58.0% of listed nanoproducts, the largest number of ENMs were

suspended in liquids (50.4%) followed by those bound to the surface (29.0%) (Figure 2.1). The other matrices account for ENMs suspended in solids (10.8%), nanostructured on surfaces (4.8%) and in bulk (2.5%), and applied in bulk forms (2.5%) (Figure 2.1). The ENMs in multiphase, nanostructured thin films and coatings comprise 0.4-0.9% of listed products (Danish Consumer Council 2015). Catalogued nanoproducts are not all-inclusive, but available data demonstrate the high application of ENMs in colloidal and surface particle matrices (Fig. 2.1).

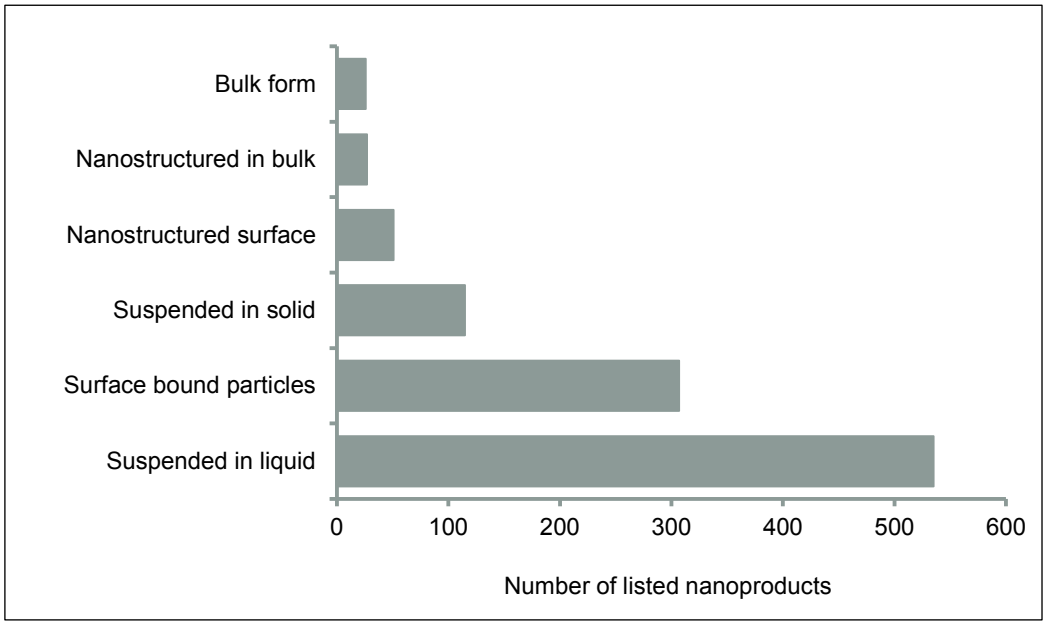


Figure 2.1 Matrix forms of ENMs components in nanoproducts.

[Source of data: (Project on Emerging Nanotechnologies 2015)].

The nAg matrices have single and multiphase assortments while other ENMs' matrices are mainly single phased. Typically, nAg is coated, embedded and dispersed as colloids or particles, while $nTiO_2$ is mainly a colloidal or particulate constituent in nanoproducts (Danish Consumer Council 2015). Also, carbonaceous ENMs are mainly embedded whereas nZnO is a colloidal and particulate ingredient, similar to $nTiO_2$ (Danish Consumer Council 2015). Thus, a large percentage of nAg and carbonaceous ENMs are likely to be disposed of to landfills while that of $nTiO_2$ and nZnO released into the aquatic environment.

In 2005, 65.0% of estimated global production of $nTiO_2$ were applied in personal care nanoproducts, sunscreens, and cosmetics (US EPA 2010a). Similarly, in 2010, 52.0% of estimated ENMs in cosmetic, coatings, paints, and pigment nanoproducts that contributed

89.0-97.0% build-up in the aquatic environment comprised $nTiO_2$ (Keller et al. 2013). These estimates place $nTiO_2$ in the lead as the most produced and released ENM in the aquatic environment.

The growth in nanotechnology and unknown quantities of ENMs released into the environment, especially aquatic ecosystems, necessitates investigating their behaviour kinetics and toxicity effects. Section 1.2 in Chapter 1 provides an overview of functional networks in the aquatic environment and likely kinetic processes that would influence them by way of ENMs' interaction, transformation, and effects. Chapter 1 introduces strategies adopted to address nanoecotoxicology issues of ENMs. The subsequent Section 2.2 describes modelling approaches used to estimate environmental mass-based concentration, distribution, and toxicity of ENMs that provide background information for this study.

2.2 Modelling in Nanoecotoxicology

2.2.1 Substance Flow Analysis Models

Substance flow analysis (SFA) are deterministic models hinged on the law of mass conservation to estimate temporal influent, matrix and sub-matrix flows, and effluent loads of chemical pollutants in the environment. They incorporate life cycle analysis within and outside matrix boundaries to assess different emission and flow pathways after the discharge of pollutants. Proposed by Baughman and Lassiter (1978), these models quantitatively evaluate risks of chemicals in the aquatic environment and have been applied to estimate environmental risks of organic and inorganic chemicals (Vighi and Calamari 1993). Studies have reported using specific principles that govern SFA modelling to estimate the flow of ENMs in environmental compartments. A review (Gottschalk et al. 2013) outlines some of the studies and their findings.

The SFA model parameters are derived using hypothetical postulation, extrapolation, and cognitive techniques (ECB 2003), which require knowledge about the production of ENMs, application in nanoproducts and life cycle emissions. Market information guides in estimating input concentrations of ENMs released from nanoproducts and extrapolating direct or indirect distributions into environmental compartments. The geospatial-temporal projections proportionally distribute quantities of ENMs from continental to regional, country, province or city boundaries using population, per capita gross domestic product, human development, and socio-economic indices. Migration and concentration of ENMs in the environment are projected using available data, presumed emission scenarios, and transfer coefficients; as

well as hypothetical estimates from surrogate environmental pollutants (O'Brien and Cummins 2010). The annual mass fluxes of ENMs account for transformation and attenuation in the environment.

Classical SFA models apply deterministic single point algorithms to estimate concentrations of ENMs (Boxall et al. 2007; Mueller and Nowack 2008; Nota 2011; Musee 2011b; Johnson and Park 2012; Keller et al. 2013; Keller and Lazareva 2014). Available input data approximated fixed-point accumulation of ENMs in environmental compartments. In contrast, probabilistic SFA models estimate likely quantities of ENMs released into the environment (Gottschalk et al. 2009; Gottschalk et al. 2010; O'Brien and Cummins 2010; Gottschalk et al. 2011; Johnson et al. 2011; Johnson et al. 2011; Money et al. 2012; Sun et al. 2014; Hendren et al. 2013; Gottschalk et al. 2015; Dumont et al. 2015). Probabilistic methods use available or imputed input data to quantify uncertainties that yield mixed spatiotemporal distributions of ENMs in environmental compartments.

Surface characteristics of ENMs were found to influence behaviour kinetics in the aquatic environment (Money et al. 2012; Hendren et al. 2013; Sun et al. 2014). For example, the stability and aggregation of ENMs are affected by adsorption properties of different surface capping agents (Hendren et al. 2013) and solubility of ionizable ENMs (Sun et al. 2014). Moreover, hydrologic factors affect migration of ENMs in rivers, which suggest potential accumulation and bioavailability in littoral and subsurface regions (Johnson et al. 2011; Johnson et al. 2011; Gottschalk et al. 2011; Gottschalk et al. 2015; Dumont et al. 2015). These findings elucidate the significance of characterizing properties ENMs and receiving environment to estimate behaviour kinetics and bioavailability.

The SFA models are large-scale and require copious data to approximate environmental concentrations of ENMs. The source, quantity, and quality of input data govern model robustness. Estimated production volumes of ENMs and knowledge on legislations related to waste and wastewater treatment in specific regions assisted in formulating assumptions and estimating environmental inflows, boundary partitioning, and outflows of ENMs. For example, high estimated production volumes of $nTiO_2$ are reflected in elevated emissions into surface water, accumulation in sediments and enhanced risks (Gottschalk et al. 2015).

The SFA models provide baseline information for assessing potential risks of ENMs in the environment since emissions are currently unknown. Few studies have used available toxicology data and projected environmental concentrations to compute the risk quotient (RQ) of ENMs (Mueller and Nowack 2008; Gottschalk et al. 2010; Musee 2011b; Nota 2011;

Money et al. 2012; Gottschalk et al. 2013). The models demonstrate varied concentrations of specific ENMs in environmental compartments, as well as high emission from effluent and accumulation in aquatic sediments.

Choice of model inputs, assumptions, and spatial-temporal variations determines the predictive outcome, comparison, and replicability of SFA models. Data paucity enhances model uncertainties, especially by applying classical modelling approaches. Assumptions of homogeneous properties of ENMs, receiving environment and emission scenarios simplify SFA models that favour mass balance estimation. These assumptions are likely to under- or overestimate concentration and risks of ENMs. Moreover, voluntary labelling of ENMs content in nanoproducts (Lomer et al., 2000) reduces the degree of certainty for SFA models that are developed using market information to estimate the concentration of ENMs released from nanoproducts into the environment.

2.2.2 Multimedia Fate and Transport Models

Multimedia fate and transport models apply multimedia box transport mechanisms to predict concentrations, mobility, transformation and accumulation of chemicals in the environment (Mackay 1979). The principles that govern SFA modelling apply in multimedia models to estimate point and diffuse emissions of pollutants. The model inputs are drawn from characterized properties of the pollutants and receiving environment (Mackay et al. 2001), with the latter having fixed conditions (MacLeod et al. 2010). The available data, assumptions and featured environmental compartments and sub-compartments define the basic rules of multimedia modelling.

Kinetic transport mechanisms determine boundary retention and trans-boundary transfers of pollutants in an environmental system under study. Multimedia models can be site-specific or generic, a classification based on availability and absence of data, respectively. Whereas site-specific models forecast fate and transport of pollutants in a specific environment, generic models theoretically simulate mechanisms replicable in any region. Adjusting model inputs permits projections in diverse exposure conditions.

Multimedia modelling involves partitioning environmental systems into several fixed and completely mixed spatial boxes, and mathematically expressing influent and effluent pollutants in each box. The net concentrations, accumulation, and persistence of contaminants are projected using influx and outflow quantities, and retention time in each box. Solubility and adsorption properties of chemicals affect their equilibrium partition in fluid

and adsorbed solid states, respectively. The partitioning coefficients and emission scenarios of pollutants influence the fate and transport processes under study. These are aspects of multimedia modelling adopted in nanoecotoxicology.

The properties of ENMs and water, residence time, and distance from discharge points affected substance transformation, migration and accumulation in river sediments (Blaser et al. 2007; Praetorius et al. 2012). Broadly, the input variables used to project fate and transport of ENMs in the water column and sediment compartments comprised information on the initial concentration of ENMs in effluent and historical hydrologic factors. The exclusion of of ENMs from air and soil as model inputs may likely underestimate outputs.

The model outputs revealed that ENMs partition as equilibrium concentration of ions, and dispersed or settled nanoparticles and aggregates (Blaser et al. 2007; Praetorius et al. 2012). The surface characteristics of exposed ENMs augment energy-controlled interactions with aquatic ligands that promote attractive attachment and aggregation, or repulsive dispersion and stability. The size and quality of attached solids depend on water velocity and total energy between interacting solids (Buffle and Leppard 1995). Hence, river flow characteristics play a significant role in fate and transport of ENMs pollutants.

Laminar or streamline flow regimes promote slow aggregation of denser and compact solids that settle faster due to their large sizes. Batch experiments have revealed slow irreversible aggregation of colloidal iron oxide ENM under static conditions (Zhang et al. 2008). Turbulent or chaotic flow regimes disperse ENMs in water columns through rapid aggregation and disaggregation effects. In sediments, drag forces acting on settled solids control the retention, transfer along the river or re-suspension of ENMs that end up collecting downstream (Blaser et al. 2007; Praetorius et al. 2012). The outcome from these models corroborate spatiotemporal build-up of ENMs in river basins estimated by SFA models (Johnson et al. 2011; Johnson et al. 2011; Gottschalk et al. 2011).

The properties of ENMs and water, meteorological and geospatial-temporal information and bioconcentration factors in aquatic biota predicted sediment deposition (Liu and Cohen 2014). Atmospheric wet depositions and ecological compartments are indirect sources of ENMs into aquatic environment, for example, nanoparticles in air trapped in rainwater, water runoff from plant leaves and trophic transfer. Notably, low inter-particle agglomeration and advection-induced migration enhance accumulation of ENMs in aquatic sediments, a phenomenon decreased by high solubility and equilibrium quantities (Liu and Cohen 2014).

The initial concentration, sediment-water partitioning coefficients, sedimentation flux, residence time and depths account for steady-state kinetic transfer and deposition of ENMs in lake water sediments (Koelmans et al. 2009). Perturbations in water-sediment boundaries, depositions in sediment compartment and decay in sediment sub-compartments affect settlement of ENMs (Koelmans et al. 2009). The external wind factors that disturb and re-suspend sediments in the water-sediment systems (Douglas and Rippey 2000) can affect low and high build-up of ENMs in shallow and deep lakes revealed by the model.

Applying modelling approaches by (Koelmans et al. 2009; Praetorius et al. 2012; Liu and Cohen 2014), an ideal SimpleBox4Nano model simulated high adsorption and hetero-aggregation of ENMs with organic matter compared to suspended solids (Meesters et al. 2014). Enhanced aggregation lowers stability of ENMs to promote settlement. The predictions assumed indirect discharge of ENMs into surface water from terrestrial and atmospheric sources, and not effluent. The exclusion of technical compartments conceivably underestimates concentrations of ENMs in effluent released into aquatic environment.

Multimedia model predictions revise fate and transport process descriptions that account for properties of ENMs and receiving water, features that lack in classical SFA model concepts. Prior characterization necessitates incorporating properties of water, pollutants, and biota as model inputs. Similar to SFA models, predictions rely on available data, emission assumptions, and distribution pathways. For example, high estimated production volumes of $nTiO_2$ translate to enhanced environmental concentrations. The model projections reveal the relevance of aquatic sediments as major reservoirs for ENMs and ligand interactions in surface alteration, fate, transport, and bioavailability of the ENMs in the aquatic environment.

2.2.3 Quantitative Structure-Activity Relationship Models

Quantitative structure-activity relationship (QSAR) models mathematically correlate structural properties of chemical compounds to physiological, chemical and biological reactivity. The QSAR models have been adopted in environmental modelling for ENMs. However, the challenges in separating inherent properties of nano and bulk forms renders the models inadequate (Alvarez et al. 2009). Principally, properties of pollutants that explain behaviour kinetics and reactivity during exposure can be used to predict the unknown reactivity of other compounds (Puzyn et al. 2011). Computational analysis of multiple descriptors and their ranking according to their strength in explaining responses assists in selection as model inputs.

Computational QSARs link thermodynamic stability of ENMs to Young's modulus rigidity measure (Toropov and Leszczynski 2006). Rigidity enhances persistence and absorbent characteristics of CNT, and bioavailability of surface adsorbed chemical pollutants (Jackson et al. 2013). Size and crystal lattice of CNT explain solubility and octanol-water partitioning coefficients (Toropov et al. 2007). The adsorption characteristics of single-walled carbon nanotubes (SWCNTs) enhance thermodynamic stability when tested alongside enthalpy of formation and size in moist conditions (Barnard 2009). Molecular connectivity, solvation, and chromatographic (solvatochromic) descriptors affect the adsorption of organic matter on multi-walled carbon nanotubes (MWCNTs) (Apul et al. 2013). Solvatochromic descriptors correlate hydrogen bonding, physicochemical and biochemical properties of molecules and charge densities to adsorption (Abraham 1993; Kamlet et al. 1983).

Experimental QSARs demonstrate significant roles played by temperature and irradiation factors in estimating toxicity. The cytotoxicity of bacteria exposed to metal and metal oxide ENMs was linked to the latter's solubility and enthalpy of formation (Puzyn et al. 2011), and electron affinity, thermal capacity, and conductivity with or without irradiation (Pathakoti et al. 2014). The two studies demonstrate the cytoxicity of ENMs' in particulate and ionic forms. At room temperature, metal oxides absorb energy to excite and transfer electrons to higher energy levels. This phenomenon is likely to enhance electron affinity and reactivity observed in non-irradiated cytotoxicity (Pathakoti et al. 2014). The level of cytotoxicity increases with increased absolute metal electronegativity and reduced absolute metal oxide electronegativity without irradiation, and reduced thermal capacity and increased conductivity with irradiation (Pathakoti et al. 2014).

The QSAR models predict linear relationships amongst multiple inputs and identified responses, which limit their application in explaining nonlinear systems. However, aspects of QSAR modelling found important to this study include comprehensive characterization of ENMs properties, selection and evaluation of multiple descriptors to identify suitable model inputs and statistical interpretation of data. Moreover, functions that associate input and response variables apply in forecasting unknown phenomena.

2.2.4 Knowledge Discovery from Data

Knowledge discovery from data (KDD) interactively and iteratively identifies patterns to extract useful information from complex databases (Fayyad et al. 1996). The KDD is a data-driven modelling approach hinged on computational intelligence and machine learning to solve problems in various fields including hydro-informatics (Solomatine et al. 2008). The

broad term of KDD covers many aspects of multivariate data analysis ranging from exploratory analysis to extraction of interesting structures in complex data sets that reveal new processes or system insights. The KDD algorithms learn from datasets with numerical and categorical values and maps relationships between distinct input and output variables.

The KDD process involves a series of steps undertaken to avail data for learning, applying learning algorithms to extract patterns from the data and interpreting the findings to support decision-making (Fig. 2.2). Data preparation steps A-B (Fig. 2.2) encompass collating domain data from relational databases, for example, laboratory assays, signals, images and text. The data collated are then aggregated, preprocessed, subsampled and transformed (Famili et al. 1997; Han et al. 2012) into organized variable arrays amenable for learning.

The key to discovering new knowledge lies within steps C-E (Fig. 2.2) that involve computational learning from data and interpretation of previously unknown data patterns or discovery of potentially unknown physical phenomena. Application of relevant algorithms depends on available data and the task. To avoid snooping bias that yields conceptually pre-determined models, Yaser et al. (2012) suggest identifying suitable learning algorithms before organizing data to be used. Data mining methods discussed in details elsewhere (Fayyad et al. 1996; Han et al. 2012) classify, cluster, predict, sequence patterns, detect variations or discover association rules in databases.

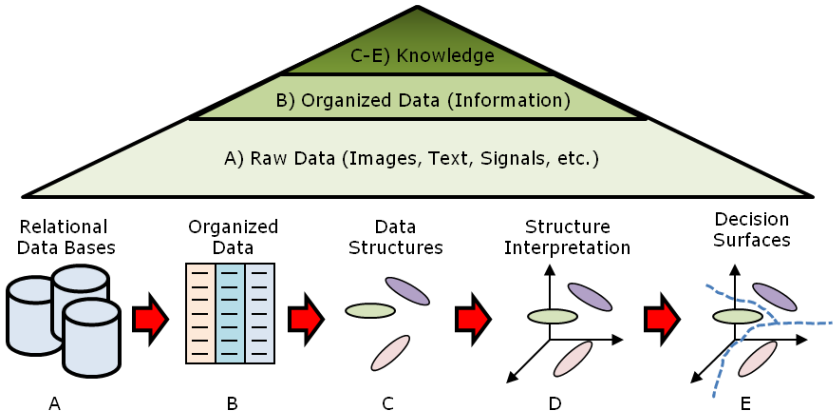


Figure 2.2 Schematic representation of the key steps in KDD process.

Adapted from: [Knowledge Pyramid (Ackoff 1989) and KDD Process (Han et al. 2012)].

The learning algorithms automatically compute linear and nonlinear patterns from data to predict or describe new quantitative and qualitative knowledge by enumerating labels of clusters, fitting models, and extracting structures. The analysis, accurate evaluation, and

novel patterns establish meaningful inference or real phenomena associated with specific variables quantified as new knowledge. The interpretation of linear or nonlinear decision boundaries extracted from data categorizes feature spaces as well as predicted or described target responses of learning problems.

Adoption of KDD as a modelling strategy in nanoecotoxicology is slow. Application of 4 supervised classification algorithms to learn from experimental data have linked cytotoxicity of fish cells to effects of physicochemical properties of ENMs and biotic factors (Liu et al., 2013). The K-Nearest Neighbour, Bagging, M5P and KStar algorithms in ABMiner tool were comparatively used to rank concentration, structural composition, surface chemistry, purity, core structure and surface charge of ENMs as important predictors of cytotoxicity (Liu et al., 2013). Another study utilized clustering and nonmetric multidimensional scaling methods to visualize susceptibility of soil bacteria to ENMs (Liu et al. 2014). Taxonomy of bacteria determines interaction levels with ENMs that can inform about terrestrial impacts.

Recently, a study applied ensemble random forest regression and classification algorithms to learn data assembled from scientific literature and predict targets that define transport of ENMs in porous media (Goldberg et al. 2015). The model effectively predicted retention fraction of nanoparticles in porous media using natural organic matter, the surface energy of ENMs and collector surfaces, and ionic strength and pH of exposure media as input variables (Goldberg et al. 2015). Retention profiles for ENMs in porous media were classified with a 60.0-70.0% accuracy score (Goldberg et al. 2015).

The KDD demonstrates potential in approximating intricate associations between various input variables, and transport and biological effects of ENMs where mathematical rules are not feasible. The ranking input variables edify their importance in estimating target responses. However, the uncertainty of KDD models lies in data available, learning algorithms selected and modelling approaches that define inputs and outputs. For instance, Goldberg et al. (2015) suggest insufficient nanoecotoxicology data is a setback in developing and validating models to enhance the understanding of ENMs' transport in porous media. Moreover, the models reported by Liu et al. (2013) accurately predicted specific biological responses depending on the algorithm used and metric-based effects targeted. Thus, the KDD models developed using nanoecotoxicology data have yielded explanatory findings, which enhance the understanding of cause and effects phenomena involving ENMs.

2.3 Summary

The advancement of nanotechnology has introduced ENMs as new pollutants released into the environment. Limited information on quantities of ENMs produced and released into the environment has necessitated the application of modelling as a tool that addresses nanoecotoxicology issues. Figure 2.3 classifies the depth and diversity of information granulation (Yao 2005) from SFA, multimedia (MM), QSAR, and KDD models discussed in Section 2.2. The information relationships depend on data content, granularity (or level of detail), and knowledge size. The modelers applied explicit knowledge in nanoecotoxicology domain from different sources to develop the models.

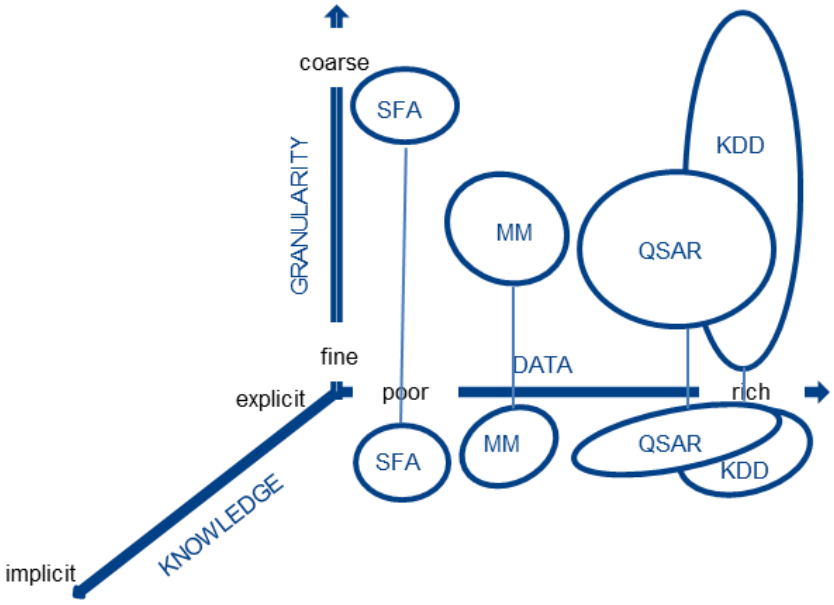


Figure 2.3 Classification of models based on data content, granularity and range of knowledge.

Adapted from: Malcolm Rider, Archestra Research[1]

The SFA and MM are large-scale models developed by applying limited available information, physical theory and first principle approaches to estimate the release, transport, and distribution of ENMs in environmental compartments and sub-compartments. The SFA models inform about priority ENMs from estimated concentrations and risks in environmental

[1] http://www.slideshare.net/MalcolmRyder/information-behaviors-versus-knowledge

matrices and are suitable for early stage risk assessment of pollutants with unknown environmental concentrations.

Conversely, MM improve the SFA modelling approach through characterization of model inputs, stochastic partitioning in environmental compartments and application of kinetic principles in forecasting fate and transport of ENMs. However, reliance on available but inadequate and uncertain data to estimate inputs, site-specific application, and diverse types and nanostructures of ENMs reduces the replicability and reliability of SFA and MM models to profile risks in environmental compartments.

The QSAR and KDD are small scale models implemented using experimental data to establish relationships between causal factors and endpoints that capture low and high levels of detail depending on data used. Mathematical correlations in QSAR models enable ranking the strength of input variables in explaining observed behaviour and toxicity of ENMs in the environment with some level of confidence. Input variables in QSAR models comprise physicochemical properties of chemical substances, in this case, ENMs, some of which are theoretically determined. However, application of predictors suitable for bulk materials limits the ability of QSAR models to explain observed phenomena influenced by ENMs.

Hypothetically, KDD models can handle imprecise data characterizing diverse ENMs and exposure conditions to extract linear or nonlinear relationships. With the capability to utilize both qualitative and quantitative data, KDD models stand out from purely deterministic SFA, multimedia, and QSAR models. The models perform on the premise of available data, whose quantity and quality determine outcomes. Nevertheless, studies that applied various KDD algorithms are few to authenticate application of these models in nanoecotoxicology. Further studies are necessary; hence, this study will apply data-driven modelling approach whose concepts are rooted in KDD.

Chapter 3 - Overview of Neural Networks and Modelling Methodology

3.1 Learning from Data using Neural Networks

Section 2.3.4 introduces knowledge discovery from data (KDD) and presents three studies that have applied specific data mining algorithms to address nanoecotoxicology issues. This Section discusses neural networks as a machine learning technique applied in this study. Neural networks are a computational intelligence tool widely applied in categorizing and approximating unknown linear and nonlinear complex phenomena whose data have predictive patterns but a theoretical understanding of existing relationships is limited (Yaser et al. 2012). They exhibit data mining and machine learning capabilities where the former discovers knowledge from data and the latter predicts and optimizes learning from examples. Both are aspects of the same discipline that often intertwine.

The architecture of neural networks was inspired by biological neurons (Rosenblatt 1958; Rosenblatt 1962), which transmit electrochemical signals via complex interconnections of nerve cells to describe the functioning of the nervous system. Nodes or neurons make up networks, often referred to as artificial neural networks, to represent abstract forms of biological neurons that imprecisely mimic how neuron transmitters process information.

Neural networks algorithms learn from diverse datasets to estimate uncertain nonlinear relationships mapping inputs to outputs (Swingler 1996). They are adaptable to learn inductively from imprecise data in complex databases, generalize functions, and develop nonlinear models. Application of neural network algorithms is suitable in many domains that present with copious data but lack theory to explain phenomena, or sufficient theory and abundant data but challenging to develop physical models (Basheer and Hajmeer 2000). Scientific literature documents application of neural networks algorithms in engineering (Haykin 1998; Aldrich 2002), ecology (Lek and Guégan 1999; Recknagel et al. 1997; Wei et al. 2001) and water resources (Maier and Dandy 2000). Application of neural networks in human nanotoxicology has been reported (Winkler et al. 2014) but not in nanoecotoxicology.

Learning algorithms can handle discrete, Boolean, interval or integer data types and learning styles categorized based on available labelled and unlabelled data as well as learning tasks,

for example, supervised, semi-supervised, unsupervised and reinforcement. Meaningful and informative classes of data define labelled data (Eqn. 3.1).

$$\mathcal{D} = \left[\left(x_n^{(j)}, y^{(j)}\right), n = 1,2,3, \ldots, N\right]_{j=1}^{m} \tag{3.1}$$

Where $\mathcal{D}$ is labelled data; $x_n^{(j)}$ represent concurrent vectors in an $m \times n$ input data matrix $(x_1, x_2, \ldots, x_N \in X)$; m is the number of training examples; $y^{(j)}$ is a finite set of values for an $m \times 1$ target or response vector, Y; and j is the *j*th location of training examples. Unlabelled data describe easily accessible non-classified data not well defined, for instance, images, and molecular structures.

Semi-supervised learning classifies and predicts outputs from both labelled and unlabelled datasets. Unsupervised learning randomly extracts unknown associations in unlabelled datasets based on commonalities. Reinforcement learning is guided probability learning that assigns scores to correct or wrong prediction while estimating labelled targets. However, the algorithm searches to find the correct output without teacher guidance. Section 2.4.4 in the previous Chapter describes the application of KDD methods to extract meaning from labelled experimental data having qualitative and quantitative variables. This study will focus on supervised learning from labelled data on behaviour kinetics and toxicity of ENMs to predict identified continuous target responses.

3.1.1 Supervised Learning

Supervised learning is a predictive approach that estimates categorical levels or classes and real values of responses in classification and regression problems, respectively. In classification, training examples of the response may be dichotomous (binary) or polytomous (> 2), denoted as $1, \ldots, C \in Y$, where C is the number of classes to predict. Notable classification problems have used data by Fisher (1936) to classify *setosa*, *versicolor* and *virginica* species of the genus *Iris* flower plant based on lengths and widths of their sepals and petals. This study focuses on regression learning, which estimates real-valued outputs compared with expected target values using regression analysis.

Supervised learning algorithms interpolate data by probabilistic iterations to converge solutions that approximate relationships and targets. The learning algorithms pair known inputs to predict known targets in labelled datasets. An optimized numerical fit of known predicted targets enables prediction of previously unknown outputs when input information

changes. Supervised regression learning models include perceptron (Rosenblatt 1958; Rosenblatt 1962), multilayer perceptron (MLP) (Rumelhart et al. 1986) and radial basis function (RBF)(Broomhead and Lowe 1988) elucidated in textbooks (Bishop 1995; Haykin 1998; Marsland 2009). The perceptron predicts linear functions and binary output components while the RBF estimates kernel distances between inputs and targets. This study will apply the MLP model that approximates functions to explain complex nonlinear relationships.

3.1.2 Multilayer Perceptron

The multilayer perceptron (MLP) model contrasts the perceptron and other statistical regression models characterized by straightforward linear mathematical functions to explain input-output relationships. The MLP structure comprises the input, hidden and output layers that learn from labelled data using feed-forward and error back-propagation signal transfer processes. Figure 3.1 graphically represents an MLP structure with connectivity between causal and effect nodes in each layer illustrating directed graphs of system dependencies. The number of nodes or neurons determine the size of each layer.

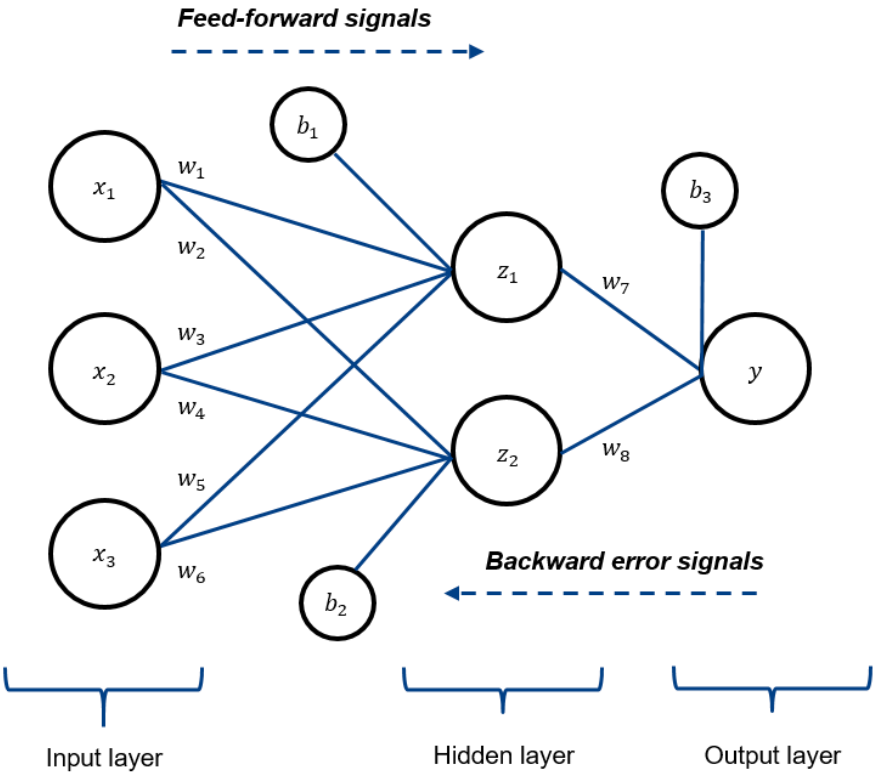


Figure 3.1 Illustration of a 3-layer MLP structure with 3-inputs, 2-hidden and 1-output nodes.

The numerical weights vector $(w_1, w_2, w_3, \ldots w_8 \in W)$ comprise scalar quantities representing signal strength of each interconnection while bias nodes, b_1, b_2 and b_3, connected to output nodes in the hidden and output layers hold constant values that control activation of predictions. The algorithm nonlinearly transforms input vectors that are in turn linearized by weight vectors in the hidden layers to map outputs as targets (Yaser et al. 2012). In the feed-forward pass, external signals of input vectors $\left[x_n^{(j)}\right]_{n=1}^{3}$ are transformed by a nonlinear activation function (f_a) and summed into nodes of hidden layer's space as $z_1^{(j)} = f_a\left(w_1^{(j)}.x_1^{(j)} + w_3^{(j)}.x_2^{(j)} + w_5^{(j)}.x_3^{(j)}\right)$ and $z_2^{(j)} = f_a\left(w_2^{(j)}.x_1^{(j)} + w_4^{(j)}.x_2^{(j)} + w_6^{(j)}.x_3^{(j)}\right)$. A second linear or nonlinear activation function (f_b) transforms summed signals in the hidden layer to estimate a signal in the output layer, $\hat{y}^{(j)} = f_b\left[\sum z_n^{(j)}\right]_{n=1}^{2}$. Equations 3.2 to 3.4 illustrate a generic learning process.

$$x_n^{(j)} \in X \xrightarrow{\Phi} z_n = \Phi\left(x_n^{(j)}\right) \in Z \ldots\ldots\ldots input \tag{3.2}$$

$$\hat{Y} = f_b(Z) \ldots\ldots\ldots output \tag{3.3}$$

The prediction for an approximate function (Eqn. 3.3) sought:

$$f: X \to \hat{Y} = f_b\left\{\sum_{k=1}^{M}\left(w_n^{(j)}.x_n^{(j)} + b_k\right)\right\} = sign\left(\vec{W}\Phi(X)\right) \tag{3.4}$$

where $\hat{Y}$ is the estimated output of target vector Y; Φ is a nonlinear transformation function; b_k is a bias term associated with nodes; and M is the number of hidden layers. The f_a and f_b notations symbolically separate activation functions in the input-to-hidden and hidden-to-output layers. Practically, similar functions may apply in both signal transfer situations. Unlike statistical models where mathematical functions of inputs and regression coefficients represent linear outputs, Eqn. 3.4 illustrates a linear output estimated by a combination of nonlinear transformations. Bias thresholds are function weights that determine activation signals of nodes, which trigger or terminate activation if the threshold function approaches set limits of outputs in hidden and output layers (Bishop 1995; 2006).

Output errors in the forward pass are propagated back to form new weights in preceding hidden and input layers to create renewed input parameters, and used once more in feed-forward iterations until outputs converge to estimate the target with minimum errors. The

MLP estimates are infinite and malleable in estimating multiple solutions to the same problem, a characteristic that generalizes learning rather finding a determinate prediction. Discussed elsewhere are comprehensive details of the MLP model (Bishop 1995; Haykin 1998; Bishop 2006; Marsland 2009; Murphy 2012) and will not be repeated in this Section.

For new data, whose trend is unknown, the design of a network structure that yields good performance requires some trial and error procedures. The design involves selecting network size, data sampling size, activation functions, learning rate and learning momentum. Network sizing involves making decisions about the number of hidden layers that offers a compromise between a good fit that accurately predicts data and applicable generalizing function that can predict outputs of unknown targets using new inputs. The number of hidden layers can be ≥ 1, by which multiple layers require fewer W to approximate functions (Flood and Kartam 1994 and Tamura and Tateishi 1997, as cited in Maier and Dandy 2000) and a single layer having adequate W can estimate any function (Hornik 1989). To apply single or multiple hidden layer(s) depends on the learning problem.

While the number of input and output neurons represents known input and target variables, the number of hidden neurons is arbitrary depending on the learning problem. The number of hidden layers and the number of neurons per layer determine the number of neurons in the hidden layer. Proposed by Hecht-Nielsen (1990), Equation 3.5 can guide the design and practical determination of a range of feasible network topologies.

$$\frac{X}{(2+1)} < h_n < (2X + 1) \tag{3.5}$$

Where X is the maximum space of input variables; and h_n is the number of hidden neurons. Electing hidden neurons less than the number of input variables increases network generalization ability while more neurons enhance accuracy (Swingler 1996; Haykin 1998). Practical application of neural networks can determine an optimum network size depending on the number of input variables in each learning problem.

Learning from labelled data informs about weights and biases that determine the strength of input variables in predicting the target. Data may be pre-partitioned into training, testing and validation sets before learning or randomly partitioned during learning. The data in the training set approximates functions by learning weights and biases. The accuracy of trained models, measured by using the validation set data, determines standalone or combined models to implement. The testing set data verifies the performance of the implemented

models by estimating points outside of the training dataset. Section 3.1.4 discusses details of data splitting techniques.

The logistic and tanh are commonly used sigmoidal activation functions that transform information signals by compressing data nonlinearly to fit within the [0,1] and [-1, 1] active transfer regions, respectively. Differentiable non-sigmoidal functions, for example, polynomial and sum of cosines, were found applicable on noiseless input data having nonlinear relationships (Maier and Dandy 2000). The identity activation function transfers information from input to output linearly. In view of output prediction from the hidden space (refer to Fig. 3.1), derivatives of activation functions are,

$$f_b(z_n) = (Z_n')|_{-\infty,+\infty} \,..\, y' = \frac{1}{\gamma} \ldots . identity \tag{3.6}$$

$$f_b(z_n) = \frac{1}{1+e^{-z}}|_{0,+1} \,..\, y' = \gamma y(1-y) \ldots . logistic \tag{3.7}$$

$$f_b(z_n) = \frac{1}{1+e^{-z}}|_{-1,+1} \ldots . y' = \gamma(1-y^2) \ldots . tanh \tag{3.8}$$

where γ is a unity derivative. Authors suggest applying different activation functions in the hidden and output layers, for instance tanh and identity (Bishop 2006; Yaser et al. 2012) and tanh and logistic (Murphy 2012), respectively. Outputs of logistic functions embody genuine probabilistic learning suitable in solving classification and logistic regression problems. The identity function emulates the perceptron by linearizing nonlinear converged inputs in the hidden layer(s) to estimate normalized or non-normalized outputs. The hyperbolic function is symmetrical around zero, and offers a broad range of numerical values applicable in both hidden and output layers to boost convergence of solutions.

Neural networks learn iteratively where the number of iterations describes the frequency of input vectors utilized in error propagation. The number of iterations needs to be large enough to promote learning, although too large a number could result in overfitting when the algorithm memorizes errors rather than learns. Moreover, too few training examples constrain complete learning that may lead to underfitting the model. For an iteration process to converge, the learning rate (η) and momentum coefficient (μ) determine the progression and stability of training tasks. Simply, the two coefficients represent the speed with which a solution for a learning problem iterates to convergence. The cost of convergence may be high or low when η and μ are too small or large, respectively. Constant or adaptive η and μ

have been used to improve optimization of specific learning problems (Basheer and Hajmeer 2000; Maier and Dandy 2000).

3.1.3 Coding and Scaling Data

The performance of neural networks depends on the structure of available datasets and the scale of input and response data. Often, continuous and categorical input variables bear dissimilar scales owing to diverse units of measurement and nominal levels, respectively. Nominal, ordinal, interval or ratio data are measurement scales that classify different variables found in datasets (Webb 2002). Nominal data contain unordered labels suitable for qualitative analyses and ordinals that are ordered but not statistically analyzable. Intervals present a range of levels with possible meaningful differences while ratios are well-defined ordered, exact and interval values that can be analyzed and inferred.

Coding and scaling (or normalization) are important data pre-processing steps that influence sensitivity of neural networks models to input variables. Coding applies to both qualitative and quantitative data but this study emphasizes the technique for the former data type and scaling for the latter data type. In both instances, coding and scaling methods aim to transform data values of inputs and targets by converting them into pliable formats for learning while preserving their original denotations. Moreover, it is common to code qualitative data for nonlinear regression or classification learning, but exceptional for regression learning.

The MLP handles numerical and binary categorical data, hence, polytomous categorical data require numerical coding for meaningful learning. Various planned contrast coding schemes have been applied in qualitative research (Wendorf 2004; Davis 2010). Coding procedures assign numerical values to different nominal levels of categorical variables to distinguish (dummy), compare (effects), correlate (orthogonal) or dissociate (non-orthogonal) them (Davis 2010). For an N number of nominal levels, dummy coding, a type of non-trend contrast scheme, assigns 1's and 0's to nominal levels to create up to $N-1$ input variables (Cohen et al. 2003), has been applied in neural networks classification learning.

Assuming A, B, C and D to be four nominal levels of a categorical variable, Table 3.1 illustrates ordinal and unary (thermometer N and $N-1$, 1-out-of-N, and 1-out-of-$N-1$) contrast coding schemes applied in pre-processing data for neural networks learning. The ordinal scheme assigns integer values $(1, 2, 3, \ldots, n)$ to nominal levels, which reduces dimensionality by creating a single variable (x_1). Unary schemes introduce potential input

variables (x_1, x_2, x_3, and x_4). Details, advantages and disadvantages of other coding schemes, namely effects, repeated, deviation, Helmert, difference, orthogonal and non-orthogonal, can be found elsewhere (Cohen et al. 2003; Wendorf 2004; Davis 2010).

Table 3.1 Illustration of coding schemes using four fictitious nominal levels of categorical variable

Categorical variable	Ordinal	Thermometer N				Thermometer N-1			1-out-of-N (N)			1-out-of-N-1 (N-1)			
Nominal level	x_1	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_1	x_2	x_3	x_4	x_1	x_2	x_3
A	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0
B	2	1	1	0	0	1	0	0	0	1	0	0	0	1	0
C	3	1	1	1	0	1	1	0	0	0	1	0	0	0	1
D	4	1	1	1	1	1	1	1	0	0	0	1	0	0	0

Ordinal and thermometer schemes were found to improve the accuracy of classification learning from data having repeated nominal levels (Fitkov-Norris et al. 2012). Still, ordinal schemes can introduce ordering of categorical levels because of numerical distances created between them while unary schemes are suitable for variables displaying fewer nominal levels. Fitkov-Norris et al. (2012) suggest that coding methods applied in neural networks are not justifiable since they provide general concepts without a clear practical guideline. Thus, selecting a coding scheme depends on the learning problem and may involve trial and error procedures requiring a good understanding of available data.

Linear and nonlinear methods are applied to rescale quantitative data to lie within signal transfer boundaries of MLP activation functions, for instance, identity $[-\infty, +\infty]$, logistic $[0, 1]$, and tanh $[-1, +1]$. Scaling enhances search and convergence of targets, which prevents the learning algorithm from being stuck at the local minima or taking a long time to converge at the global minima. Moreover, scaling improves biases on error surfaces and weight decay, with high errors exhibited by data having large variances. Considering an input variable, $x^{(i)}$, minimum-maximum normalization (Eqn. 3.9) and standardization (Eqn. 3.10) are common scaling techniques reported (Swingler 1996; Han et al. 2012).

$$x^{(j)\prime} = \frac{\left(x^{(j)} - \mathrm{x}_{min}^{(j)}\right)}{\left(\mathrm{x}_{max}^{(j)} - \mathrm{x}_{min}^{(j)}\right)} \tag{3.9}$$

$$x^{(j)\prime} = \frac{\left(x^{(j)} - \bar{x}\right)}{\sigma} \tag{3.10}$$

where $x^{(j)\prime}$ is the scaled value; $x^{(j)}$, $\mathrm{x}_{min}^{(j)}$, $\mathrm{x}_{max}^{(j)}$, $\bar{x}$ and σ are the original, minimum and maximum data, and the mean and standard deviation, respectively; and j is the *j*th location of training examples. Similar scaling methods apply to normalizing $y^{(j)}$ values.

3.1.4 Training and Minimizing Errors

The aim of training is to approximate an unknown target function, $Y = f(X)$. A hypothesis function, $h(X) = f(X)$ is set and a learning algorithm used to predict outputs conditioned on inputs, $h: X \rightarrow Y$. The algorithm estimates a learning hypothesis, $g(X) = f(X)$ by predicting the target function with minimal errors. The error function, broadly defined by $E\big(h(X), f(X)\big)$, measures how well the hypothesis function estimates the target function. Implementing training involves priming unrestricted parameters, such as weights, and repeatedly updating their values during line or trust region iterative search and convergence at a global minimum. Line and trust region iterations are described in depth in (Moré and Sorensen 1983; Necedal and Wright 2006). Figure 3.2 illustrates components of learning from labelled datasets.

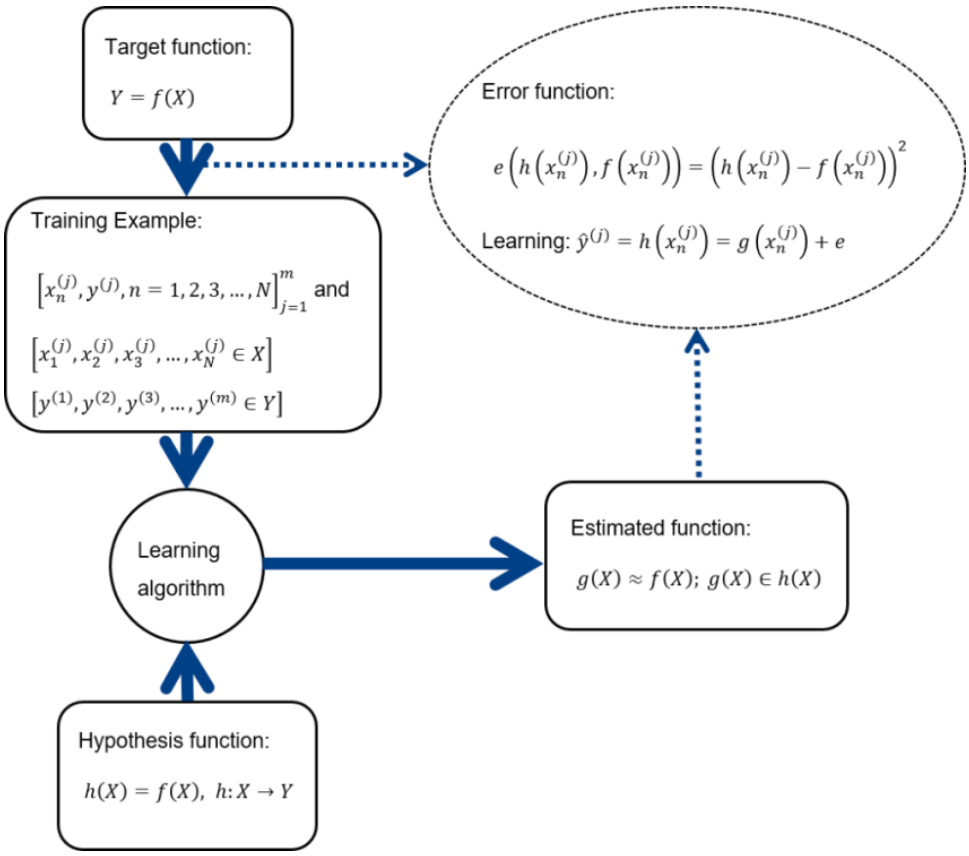


Figure 3. 2 Illustration of learning components.

[Adapted from Yaser et al. (2012)]

a) Optimization Algorithms

There are diverse learning algorithms to choose from when implementing an MLP neural networks model distinguished by iterative search and optimization methods. The gradient descent (GD), conjugate gradient (CG) and Newton methods, as well as their variants / blends, are common optimization algorithms reported in the published literature. The first order GD, and second order CG / Newton methods require low and large memory spaces, respectively (Battiti 1992), the latter caused by large computational storage of the Hessian matrix. The algorithms optimize by finding minimum data of nonlinear functions at domains with zero derivatives (Fig. 3.3).

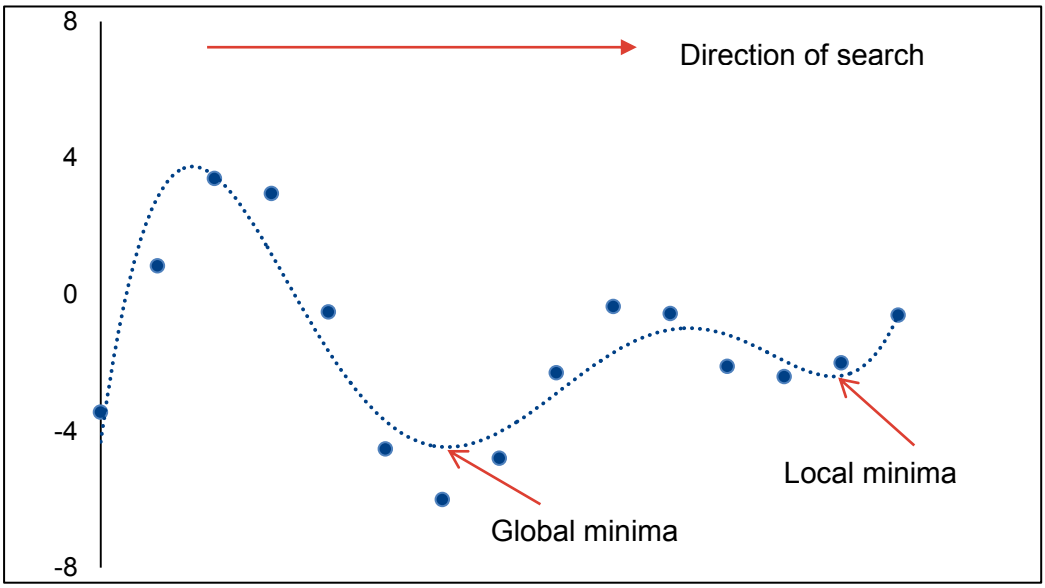


Figure 3. 3 Illustration of global and local minima using fictitious data.

The classical GD is attributed to the work of Cauchy (1847) as cited in (Petrova and Solov'ev 1997) and a common error back-propagation method (Rumelhart et al. 1986) that steeply updates parameters in the negative search direction of the function being approximated. The batch or mini-batch transfer of information in the network makes GD a slow method that produces large variances by oscillating at the minimum when m is large. The Stochastic gradient descent (SGD) (Nemirovski and Yudin (1978) cited in (Murphy 2012)) is a faster variant of GD that updates parameters in an on-line mode with every iteration. However, the SGD can overshoot the minima and generate high variances when η is large, thus, preference is given to GD (Battiti 1992).

The CG is a linear optimizer (Hestenes and Stiefel 1952) extended by (Fletcher and Reeves 1964) for nonlinear application and has been improved over time (Moller 1993; Hager and Zhang 2006) to enhance the learning rate. The CG is an inexact Newton method that utilizes modest memory space to solve linear and nonlinear regression and classification problems in large datasets. Moreover, the Levenberg-Marquardt (LM) (Levenberg 1944; Marquardt 1963) and Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Broyden 1970; Fletcher 1970; Goldfarb 1970; Shanno 1970) are common Newton optimization algorithms.

The LM blends the GD and Gauss-Newton procedures suitable for minimizing nonlinear least square regression functions. When η is small or large, the LM performs as a GD or Gauss-Newton algorithm, respectively (Hagan et al. 2014). The LM algorithm converges at local minima when the Jacobian matrix precede the Hessian matrix, which makes it comparatively fast but problematic to apply on data that is not well understood. Thus, the LM aptly solves medium to complex accurate functions having fewer parameters to optimize.

The BFGS is a popular non-constrained quasi-Newton algorithm numerical optimizer that circumvents computing the Hessian matrix by using simple matrix-vector multiplication in the search direction. The bypassing of the Hessian matrix results in super-linear convergence of solutions unlike quadratic convergence of Newton methods, such as the LM that utilizes less memory space as long as the η satisfies Wolfe conditions (Necedal and Wright 2006).

Comparatively, the LM and BFGS demonstrate fewer iteration steps to convergence than the GD and variants of CG (Burney et al. 2007). The GD is suitable in solving problems that require slow convergence to avoid under-fitting but appears to be affected by poor data scaling unlike the Newton methods that show stability (Necedal and Wright 2006). Whereas the CG is adaptable but slow, the LM and BFGS perform better in regression than classification learning of medium to complex problems. Burney et al. (2007) reported a *ca.* 2.14-fold fewer iteration steps of LM than the BFGS, although computation time increases with increasing network complexity for both methods. Further reading on learning algorithms and their variants, iteration procedures, comparative performance, application and limitations can be found in (Bishop 1995; Petrova and Solov'ev 1997; Necedal and Wright 2006; Hager and Zhang 2006; Burney et al. 2007; Murphy 2012; Han et al. 2012; Hagan et al. 2014).

b) Data Splitting for Learning

An optimization algorithm can learn from all training examples (m) in a labelled dataset $(\mathcal{D})$ but error functions would be too optimistic for trained models to generalize on new data.

However, bias-variance compromise that yield accurate and quality models to generalize well on new data involves making decisions about sizing the network, resampling data, and regularizing errors to minimize likely over- or under-fitting. The bias measures how well a model fits observed data compared to the hypothesis function while variance measures the quality of the fit. Often times domain data are insufficient, and their partitioning before or during training helps to check model performance (refer to Section 3.1.2).

Section 3.1.2 introduces network sizing additionally discussed in Section 3.2.2. The published literature reports Bootstrap (Efron 1979) and cross-validation (Stone 1974) as statistical techniques commonly used to partition m into training, validation and testing sets for training and checking models (refer to Section 3.1.2). The Bootstrap procedure randomly decomposes $\mathcal{D}$ into a training set with a probability of 0.623 sampled with replacement. Thus, the probability of data falling into the testing set is 0.368. The likelihood of values resampled more than once or not at all into any of the two sets is high. Thus, bootstrapping suitably applies to small datasets to create arbitrary random subsets that enhance learning. Additional information on bootstrap can be found in (Davison and Hinkley 1997).

The leave-one-out, K-fold, and random subsampling are common cross-validation techniques used for data splitting (Figure 3.4).

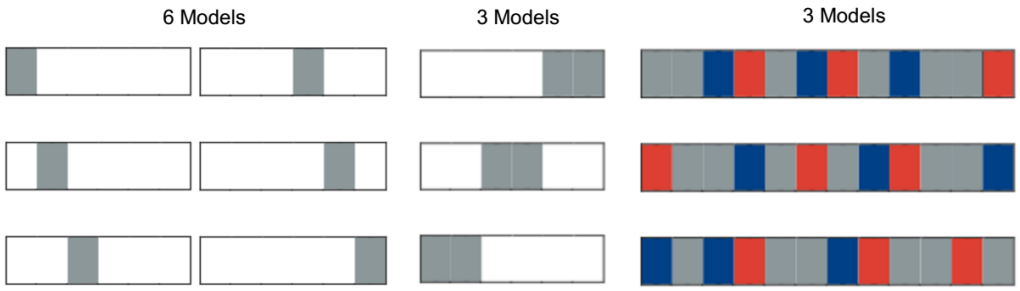


Figure 3. 4 Illustration of leave-out-one (*Left*), K-fold (*Centre*) and random subsampling (*Right*) cross-validation methods.

In repeated cycles, the training and validation sets apply in training, evaluating performance, averaging error functions and selecting models to implement. The testing set date independently evaluates model generalization by using new inputs in the implemented model to predict outputs out-of-sample. The leave-one-out method suits data paucity situations where models are trained using $m - 1$ examples and error functions tested using 1 example (Fig. 3.4 Left). The K-fold method splits data into K equal subsamples for training

and validation / testing (Fig. 3.4 Centre). Yaser et al. (2012) suggest $K = m/10$ subsamples as sufficient in enhancing learning.

The random subsampling technique randomly splits data in fixed proportions to fall in any of the three subgroups (Fig 3.4 Right). Model development may involve parallel training, validation, and testing performed in one run or parallel training and validation performed in one run followed by out-of-sample testing of selected model(s). A review by Basheer and Hajmeer (2000) outlines author suggestions of ≥ 60.0%, ≤ 30.0%, and ≥ 10.0% grouping of data into training, testing, and validation sets, respectively. As a rule of thumb, Yaser et al. (2012) suggest an 80.0% and 20.0% data split into training and testing sets, respectively. The discretion of the modeler in making a valid choice of allotting data into any subgroup depends on the quality and quantity of data at hand and expected optimal performance.

The cross-validation methods have yielded models that generalize well on new data (Bishop 1995). However, challenges associated with limited input data, complexity of measured parameters, optimistic estimated functions, and computational costs of building multiple models require statistical and probabilistic evaluations of models from training sets (Bishop, 2006; Murphy, 2012) not considered in this study.

c) Estimation of Errors

Estimated point (e) and network errors quantify algorithm capability in predicting targets. A minimized value of $\hat{y}^{(j)}$ would lie within the neighbourhood of the target value $y^{(i)}$. Thus, e is an error signal between the target $\left(y^{(j)}\right)$ and output $\left(\hat{y}^{(j)}\right)$ values (Eqn. 3.11). The network errors relate the estimate $\left(\hat{Y}\right)$ and the target (Y) functions, a measure of bias and variance defined by mean square errors (MSE) (Eqn. 3.12). The magnitude of errors determine a good estimate of observed data or an optimized function (Haykin 1998).

$$e\left(h\left(x_n^{(j)}\right), f\left(x_n^{(j)}\right)\right) = \left(h\left(x_n^{(j)}\right) - f\left(x_n^{(j)}\right)\right)^2 \tag{3.11}$$

$$MSE = \frac{1}{m}\sum_{j=1}^{m}\left(h\left(x_n^{(j)}\right) - f\left(x_n^{(j)}\right)\right)^2 \tag{3.12}$$

The quality of data and subsampling methods that split data into training and validation sets influences network errors. The outputs converge when testing errors begin to rise (Basheer and Hajmeer 2000). The error magnitude reduces with increasing h_n (Eqn. 3.5) or iteration

cycles (Basheer and Hajmeer 2000). The activation functions applied distinguish error derivatives for each weight. Output errors can be normalized by a weight decay constant to improve network performance (Eqn. 3.13) (Bishop 1995).

$$E = MSE + \frac{\alpha}{2} W^T W \tag{3.13}$$

where α is a constant value that represents weight decay, a regularization term that causes initialized weights to reduce gradually to zero. The MLP model adjusts error functions by modifying weights and summing up errors in the entire network as outlined in Section 3.1.2.

3.2 Ensemble Neural Network Model used in this Study

This study will apply ensemble supervised neural networks learning paradigm. This Section describes the methods undertaken to investigate relationships between predictor and response variables in a heterogeneous dataset. Ensemble neural networks encompass training several models and averaging them to improve output performance. Random training in neural networks yields infinite solutions to a problem. Selecting one model limits available solutions owing to learning errors and uncertainties associated with algorithms used and quality of available data. Combining finite solutions improves prediction performance that permits an understanding of complex interactions and strength of associations amongst influencing variables and target responses.

Ensemble learning involves aggregating models trained using multiple algorithms for comparison and prediction, or a single algorithm for prediction purposes. Ensembles have been found to improve the accuracy of standalone models, although aggregation may also degrade good learners (Breinan 1996). Various ensemble methods have been proposed, for instance stacking (Wolpert 1992), mixed experts (Jacobs and Jordan 1991), Boosting (Schapire 1990, 1999; Freund and Schapire 1997) and Bootstrap aggregation (Bagging) (Breinan 1996). This study focuses on aspects of the Bagging method and the references provide additional information about the other techniques.

The Bagging method has been used to build and combine multiple model versions to level their independent weakness and strengthen prediction ability (Breinan 1996). For instance, considering $\mathcal{D}$ (Eqn. 3.1) and notations used in Sections 3.1.2 and 3.1.4, the learning hypothesis, $g(X)$, minimizes $f(X)$ to approximate $\hat{Y}$ in the neighbourhood of Y (refer to Section 3.1.4) to generate a standalone model, $\mathcal{p}$. Using Bootstrap methods discussed in

Section 3.1.4, data subsamples ($\mathcal{D}_s, s = 1, 2, 3, \ldots, \mathbb{N}$), are used to build multiple models ($\mathcal{p}_s$), where s denotes independent and identical population of training examples transferred through the network to estimate diverse functions $g_s(X) \approx f_s(X)$. In regression, simple linear aggregation of $\mathcal{p}_s$ moderates their independent error functions to yield an average learning function, $g_{\mathcal{A}}(X) \approx f(X)$, and a model, $\mathcal{p}_{\mathcal{A}}$, over the entire $\mathcal{D}$, where $\mathcal{A}$ signifies aggregation. The average function would be,

$$g_{\mathcal{A}}(X) \approx f(X) = \frac{1}{\mathbb{N}} \sum_{s=1}^{\mathbb{N}} \left(f_s(X)\right) \tag{3.14}$$

Aggregation stabilizes standalone models by reducing variances at constant bias (Friedman 1997) but Bagging has been found ineffective on variances in some learning problems (Brown 2004). Thus, Bagging is suitable for heuristic models that generate diverse solutions when inputs are marginally changed, such as neural networks (Breiman 1996). Naftaly et al. (1997) showed that varying initial parameters, such as connection weights, weight decay, and learning rate as well as reducing input variables introduces variances in learners that improve with aggregation. Thus, leaving out Bootstrap resampling aspect described in Section 3.1.4 and focusing on aggregation concepts of Bagging, this study will ensemble learners trained by a single algorithm without pruning. Varying the initial random seeds will create different data subsamples and parallel averaging will create aggregated ensembles.

3.2.1 Data Collation and Preparation for Learning

Data on behaviour kinetics and toxicity of engineered nanomaterials (ENMs) generated under aqueous conditions will be collated from accessible published scientific literature for use in this study. A review of existing guidelines for collating secondary data will assist to develop criteria for extracting and compiling relevant information in this study. The model inputs will be identified from a list of physicochemical properties of ENMs, and environmental abiotic and biotic factors. The target responses will be identified from quantified adsorption and aggregation endpoints of ENMs, and their toxicity to algae and water flea, *Daphnia magna*. A structured labelled dataset for learning will be constructed holding data for identified model inputs and outputs.

Factors that influence natural phenomena, for instance, aquatic environment, can be empirically determined or described under simulated laboratory conditions. Chapter 1 introduces uncertainties of simulating the behaviour and toxicity of ENMs in an aqueous environment with the potential generation of mixed quantitative and qualitative data. In

anticipation of collating analogous data during this study, methods described in Section 3.1.3 will be used to transform the data to facilitate learning using the MLP neural network model.

Coding schemes discussed in Section 3.1.3 may be problematic to implement if categorical data exhibit many correlated nominal levels having dissimilar frequencies. This study will apply a coding scheme aimed at reducing dimensionality and potential ordering of transformed data. Frequency-based coding, a technique applied by Wilcoxon (1945) to rank and compare quantitative endpoints generated from paired and unpaired experiments, appears favourable. Using Wilcoxon (1945) method as a basis, coding nominal levels of categorical input variables will involve the following,

- Structure the learning dataset as a matrix of m training examples.
- Assign ordinal ranks to nominal levels from 1-m.
- Group similar nominal levels and replace tied ranks with mean values.
- Divide the ranks computed in bullet (3) and (4) above by the number of training examples to obtain frequency ratios (or fraction codes) between zero and one.
- Assign the fraction codes as quantitative scores to distinguish each nominal level. Table 3.2 illustrates the coding procedure using fictitious data for demonstration purposes.

Table 3.2 An illustration of frequency-based coding of categorical data

Categorical variable	Ordinal Rank	Frequency Rank	Fraction code
A	1	2.5	0.227
A	2	2.5	0.227
A	3	2.5	0.227
A	4	2.5	0.227
B	5	6.0	0.545
B	6	6.0	0.545
B	7	6.0	0.545
C	8	8.5	0.772
C	9	8.5	0.772
D	10	10.5	0.955
D	11	10.5	0.955

Key: Frequency for A = 4, B = 3, C = 2, and D = 2. Total number of training examples = 11.

For reasons described in Sections 3.1.2 and 3.1.3, nonlinear hyperbolic tanh and linear identity will serve as input and output activation functions during training, respectively. The pre-determined transfer functions determine data rescaling methods, which is the minimum-maximum normalization technique for continuous data to lie between -0.9 and 0.9 appropriate for tanh weights transfer region (Eqn. 3.15).

$$x^{(j)\prime} = -0.9 + \left(0.9 - (-0.9)\right) \times \frac{\left(x^{(j)} - \mathrm{x}_{min}^{(j)}\right)}{\left(\mathrm{x}_{max}^{(j)} - \mathrm{x}_{min}^{(j)}\right)} \tag{3.15}$$

Findings by Fitkov-Norris et al. (2012) demonstrated that coding nominal levels influenced accuracy more than rescaling inputs. Thus, quantitative scores assigned to nominal levels of categorical data will remain unaltered. However, application of standardization method will scale the response data to attain zero mean and unity standard deviation (Eqn. 3.10). Normalization of input data, as opposed to their standardization, can prevent prospective generation of values that fall outside the region of tanh activation function.

3.2.2 Network Design

By way of a known number of inputs and targets identified from collated information, the number of hidden neurons in the hidden layer will determine the network configurations. Sections 3.1.2 and 3.1.4 describe the significance of connection weights (W) and hidden neurons (h_n) in transferring and modifying information signals through the network. The current published literature does not provide specific design concepts for determining the number of hidden layers, which necessitates a problem-based approach that involves trial and error procedures based on the learning problem.

Previously, network structures have been determined by fixing the number of hidden layers and optimizing h_n or fixing h_n and varying W to optimize the topology (Maier and Dandy 2000). This study will adopt the first approach and fix a three-layer network topology having one hidden layer (refer to Fig. 3.1 in Section 3.1.2) to be used in all experiments. Theoretically, Hornik (1989) and Cybenko (1989) have shown that single hidden layer MLP neural networks models can approximate functions of any learning problem. Practically, Lek et al. (1996) used ecological inputs to estimate the abundance of fish using a single hidden layer MLP neural networks model. Thus, one hidden layer network appears to be sufficient for use in this study, whose structural connectivity will be practically determined.

The optimum h_n will be practically determined through training and selecting models based on minimum error functions. Equation 3.5 in Section 3.1.2 will be used to compute feasible h_n within the minimum and maximum range depending on the number of input variables. The minimum and maximum h_n will be computed using the first part of Eqn. 3.5 and subtracting one from the number of input variables, respectively. A regularization parameter (α) (Eqn. 3.13) will be practically determined through trial and error procedures based on available data.

3.2.3 Training Ensemble of Neural Networks

The model will be implemented using neural networks data mining tool in Dell STATISTICA 64-bit Software Version 13 on a Microsoft Windows 7 Intel Core i3 PC platform. The tool displays ease of navigation with a user-friendly icon graphical user interface and data uploading workspaces. It also integrates automated and customized modelling options built-in the GD, CG, and BFGS optimization algorithms described in Section 3.1.4 and display ease of model deployment, analysis, and custom predictions. The complexity and accuracy of learning problems in this study are unknown. Thus, the BFGS algorithm appears suitable for use in this study because of better performance in regression learning from small to medium datasets, faster and numerically stable linear convergence of estimated functions, self-correction and data scaling invariant property (Necedal and Wright 2006).

The MLP neural networks model development will involve concurrent assessment, approximation, and analysis stages. The assessment stage will establish optimal network configurations based on the h_n and regularization parameter (α) that yield models approximating target functions with minimal errors. By deploying models retained during the assessment and approximation stages, the analysis stage will involve the following procedures

- Compare, nominate, and combine trained models to create ensembles.
- Determine the relative importance of input variables in predicting the target responses.
- Determine the model sensitivity to changes in input variables to establish likely relationship trends.

Applying the cross-validation method, compiled datasets will be randomly partitioned into 90.0% and 10.0% training and testing subgroups (or sets), respectively. Using five initial integer values or seeds varied between 1 and 10,000, further random partitioning of the

training set data will create parallel a training (80.0%) and validation (20.0%) set during model building exercises. Data partitioning will be similar for different network topologies assessed by varying h_n. Varying random seeds is not a foolproof technique since randomization may repeatedly place the same data values in the same set. Nevertheless, random subsampling will yield different model outputs built using different populations of data and retaining those subsequently analyzed. Figure 3.5 graphically illustrates procedures of data splitting and regression learning.

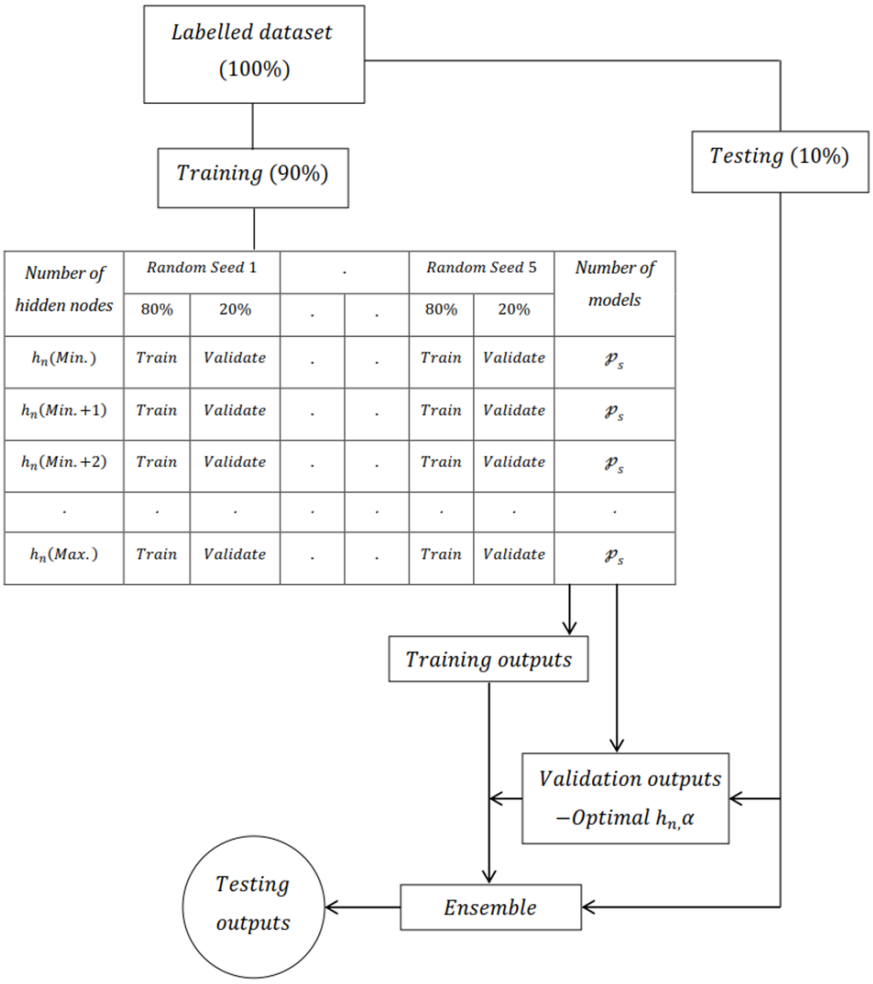


Figure 3. 5 A graphical illustration of regression learning procedures.

To aggregate models, authors suggest evaluating the quality of validation data subsets (Haykin 1998) and variances from the sum of square errors (Hastie et al., 2008). In both approaches, low-performance errors yield suitable models. To select an optimal network design as well as standalone models from those retained to aggregate, the following procedures will be applied,

- Vary the α value within suitable regions and determine validation set MSE of retained standalone models.
- Average the validation set MSE over the number of retained models for every α value tested. Use the holdout data to test generalization of standalone and aggregated models based on individual α values.
- Select the optimal α based on lowest average validation set MSE.
- Utilize the selected α value to train standalone models while varying number of hidden neurons.
- Select the optimal network topology based on the average validation set MSE.
- Aggregate models based on optimal network topology.
- Use the testing set to assess generalization ability of fitted ensemble models using the 90.0% training and validation set.
- Perform regression analysis of the ensembles and estimate normality using Shapiro-Wilk analysis (Shapiro and Wilk 1965).

The topology of selected models for ensemble prediction will depict the number of input variables as input nodes in the first layer, optimal h_n in the second hidden layer and target response(s) as output node(s) in the third layer.

3.2.4 Importance of Variables

The performance of neural networks model depends on various combinations of input variables to estimate solutions. Combining different inputs generates varying errors of approximated functions as well as magnitudes of outputs. For instance, where N is the number of input variables, possible combination of solutions is would be 2^N. The best solution is that which yields the least errors. Model inputs drawn from properties of ENMs and aqueous environment can change and have error implications on approximated target responses. Thus, performing sensitivity analysis (SA) measures the uncertainty of input variables for different purposes such as enhancing the understanding of systems being modelled and outlining recommendations for decision making. Yeung et al. (2010) highlights the application of SA in neural networks modelling.

In published literature, methods used to determine variable importance include partial derivatives (Dimopoulos et al. 1999), perturbation (Scardi and Harding 1999), 'Profile'

(Gevrey et al. 2003) technique proposed by Lek et al. (1996), weights (Garson (1991) as cited in (Gevrey et al. 2003; Goh 1995), and forward and backward stepwise (Gevrey et al. 2003) and their variants. Comparative performance of these methods has been reported using simulated data (Olden et al. 2004; Dedecker et al. 2005) and more information on applications can be obtained from the given references. In this study, SA of aggregated models will identify relative importance of input variables in predicting the targets and establish explanatory contribution of each variable to the prediction relationships to recommend future considerations in nanoecotoxicology issues.

The algorithm that will be used iteratively determines variable sensitivities by evaluating network errors in a finite manner, replacing values of each input vector with its mean and estimating errors of predicted functions using substituted vectors. Different input variable combinations that yield different sensitivities are likely. However, this study will not utilize outputs of inbuilt SA because sensitivities of qualitative data will also be determined. Instead, a perturbation method proposed by Potts (1999) as cited in (Francis 2001) and the 'Profile' technique (Lek et al. 1996) will be used to determine the relative importance of input variables and model response to changes in the variables, respectively.

The concepts of the two methods are founded on principles of partial derivatives, whereby a specific variable is differentiated with respect to other variables of a given hypothesis (or function) (refer to Fig. 3.2 for clarity). Initially, data will be statistically classified to lie within the minimum-maximum range. The first SA method (Francis 2001) will be implemented using the following procedure,

- Deploy the fitted ensemble neural networks model(s) to display approximated outputs using the 90.0% training and validation data sets, and ensemble MSE.
- Fit new models by utilizing all training examples based on the following custom data,
    - For continuous data, replace values of each input variable using the mean value while holding other variables constant.
    - For categorical data, replace all the numerically quantitative scores of each categorical variable while holding other variables constant.
- Compute the MSE of newly fitted models to define the relative importance of each input variable by comparing with the MSE of the full model in (bullet 1) above.

- In descending order, rank the variables by assigning high importance to those that show high MSE's when varied, and low importance otherwise. Moreover, rank each nominal value of categorical variables to show relative importance of their individual contribution.

Contrasting the first method whereby all training examples will be applied, the following steps will be followed to perform the second SA (Lek et al. 1996),

- Statistically, analyze the data to describe the minimum and maximum range of values for each input variable.
- Compute sequential steps by subtracting the minimum from the maximum data and dividing by 11 to generate 12 equal interval values of each continuous input variable over the minimum-maximum range.
- Deploy the fitted ensemble neural networks model(s) to display approximated outputs using the 90.0% training and validation data sets, and ensemble MSE.
- Fit new models based on the following custom data,
    - For continuous data, replace values of each input variable using the 12 simulated values while holding other variables constant at their minimum, lower quartile, median, upper quartile and maximum data.
    - For categorical data, successfully replace all the numerically coded values of each nominal level while holding other variables constant at their minimum, lower quartile, median, upper quartile and maximum data.
- For continuous data, plot the input versus predicted response profiles to display relationship trends.
- For categorical data, plot the input versus predicted response histograms to show the contribution of each nominal level to the response.

The comparison of the simulated outputs from the two sensitivity analyses will form the content of interpretations and discussions.

3.3 Summary

Ensemble modelling paradigm pursued in this study will apply multilayer perceptron (MLP) neural networks regression algorithm to predict responses identified from behaviour kinetics and toxicity of $nTiO_2$. The prediction will permit the understanding of complex interactions and strength of influencing variables in estimating the responses. A considerable amount of

data is desirable for neural network models to be useful. The modelling approach will involve the following stages; (i) gathering accessible data from scientific reports on behaviour kinetics and toxicity of $nTiO_2$, (ii) preparing the data for modelling, (iii) training models using MLP algorithm, and, (iv) testing performance, accuracy and sensitivity of models (Fig. 3.6). Circular arrows indicate phases involving recurrent steps.

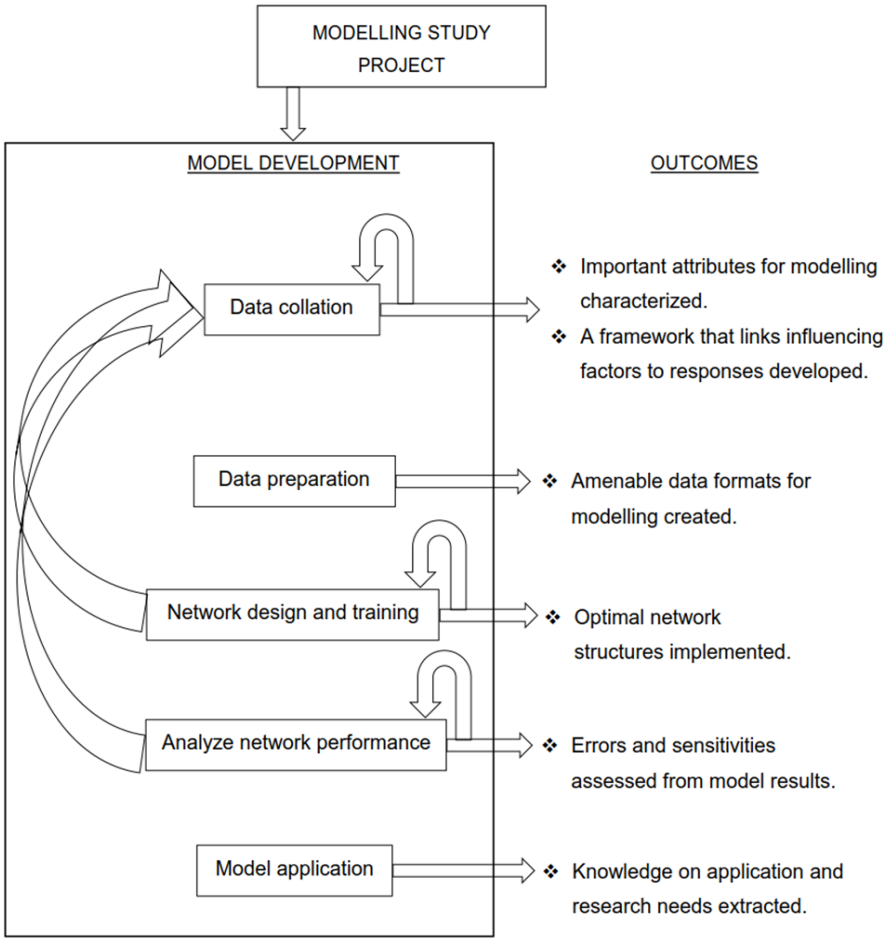


Figure 3.6 Steps in executing the modelling study.

[Adapted from Swingler (1996)]

Chapter 4 - Developing a Dataset for Learning

4.1 Secondary Data and Current Databases

Chapters 1 and 3 introduce neural networks as a data-driven modelling technique. Thus, collation of data related to behaviour kinetics and toxicity of ENMs from accessible sources is an important step in this research study. The absence of data makes decision-making a challenging process. Sections 3.4.1 and 3.4.2 of Chapter 2 highlight how inadequate data necessitates the application of theoretical notions and relevant principles to actualize systems not well understood as a basis for decision-making. Available data may have levels of uncertainty because of their derivation methods but are usable to estimate actual phenomena. Data-driven models can bridge knowledge gaps by predicting processes, associations and systems whose data may be difficult to measure.

Experimental designs, test protocols, and statistical analyses used to generate primary data are procedures grounded on fundamental scientific principles that support data-driven modelling. Hence, scientific publications on behaviour kinetics and toxicity of ENMs provide a source of readily available secondary data. Accessing and applying secondary data in predictive modelling saves time and costs since large samples of representative data are available from multiple sources. Moreover, secondary data facilitates application of multiple variables that would be difficult to investigate in a single experiment. However, the process of obtaining secondary data may be costly depending on the quality of accessible information, and objectives of the intended study.

A protocol that guides information sharing amongst different research communities on nanoecotoxicology is lacking. When this study commenced in 2011, an evaluation of accessible databases could not establish potential excerption of relevant information on variables related to behaviour kinetics and toxicity without rigorous data collation procedures. Records showed limited databases used as platforms to share information on nanoecotoxicology despite growing literature on the subject. Based on personal observations, accessible databases are public, private, or exclusive. Public databases are completely free to users; private ones require login registration and are partially accessible while exclusive ones are pseudo-open and accessible to specific groups, for example, researchers. Appendix 4.1 illustrates notable databases.

Open access databases linked to nanotechnology store unstructured information on production, manufacture, application, and commercialization of ENMs as well as research protocols and articles among others. They include public inventories, which catalogue nanoproducts based on information that is crowd-sourced or voluntarily provided by manufacturers. Information from these databases can be retrieved, reused, and referenced. However, the authenticity of information for listed nanoproducts, which exhibit incomplete description of ENMs' content, rests with the providers and would be challenging to corroborate. Other inventories provide detailed information on properties of ENMs and exposure systems that are useful in this study; however, their registries do not list ENMs that are potential case study materials for this dissertation.

Accessible business-oriented databases are proprietary in nature. They provide useful insights on broad aspects related to the growth of nanotechnology from small-scale industry to the global arena. However, these databases provide little useful information on potential exposure mechanisms and toxicity of ENMs for excerption. Researchers have developed web-browser interfaces to facilitate online access to relevant research articles and other published information and construct in-house repositories. The registries support institutional model development linking ENMs to biological effects. Restrictions on stored information require login registration and approval, which is a lengthy process and often unsuccessful.

There were limitations to retrievable information from repositories. First, quality of information held in public repositories depends on the sources, some of which excluded details considered beneficial in this study. Second, access to desired information was restricted due to protracted trade legislations and agreements. In this regard, the study designed to collate secondary data usable for data learning.

4.2 Conceptual Modelling Framework

The first objective, which is to "Establish inherent physicochemical properties of ENMs, and environmental abiotic and biotic factors influencing behaviour kinetics and biological toxicity during exposure" guides a conceptual modelling framework developed for this study (Fig. 4.1). The framework outlines the behaviour, fate, and biological effects of ENMs after emission into aquatic ecosystems. Directed graphs link influential factors to target responses in the direction of the arrows. From a series of steps presented in this Chapter, the first objective answers the research question, "What inherent physicochemical properties of ENMs and abiotic factors influence their behaviour kinetics in the aquatic environment?"

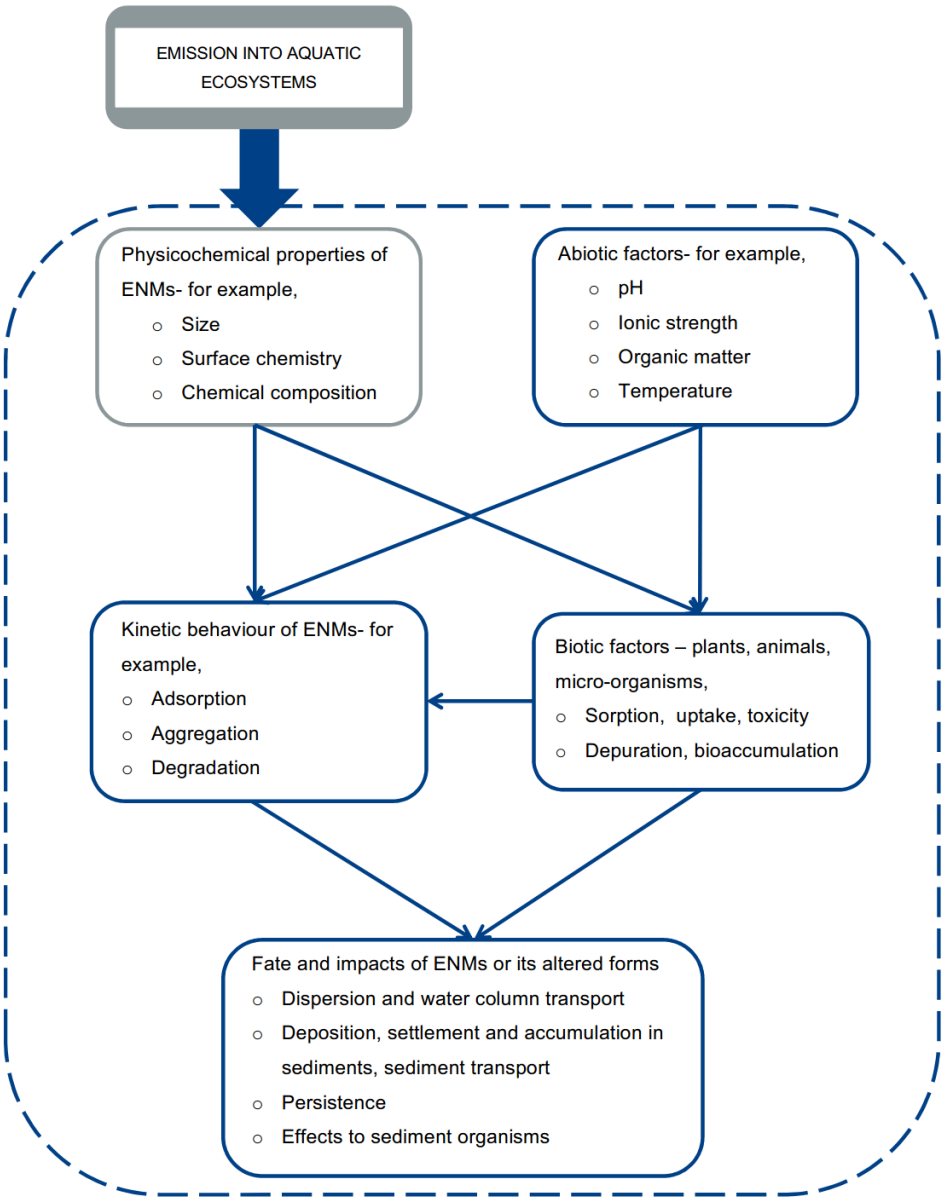


Figure 4.1 Generalized conceptual modelling framework illustrating causal and effect relationships involving ENMs in the aquatic environment.

Understanding the properties of ENMs and kinetic processes that influence mobility, interaction, and effects after their release into aquatic environment enlighten about relevant data to collate from domain scientific literature. Properties of ENMs determine the nature of their thermo-, chemo-, and photolytic transformation under favourable environmental conditions. Kinetic processes govern the state of ENMs, for example, as released or environmentally altered, in which they disperse and migrate, interact with organisms, deposit and settle in sediments, or accumulate and persist in ecosystems. Organisms are a biological matter that interrelates with ENMs as either food or adsorbing surfaces, leading to

uptake, bioaccumulation, deposition, settlement as aggregates and toxicity impacts. Subsequent Sections present data collation and analyses procedures, factors influencing behaviour kinetics and toxicity of ENMs and construction of a learning dataset.

4.2.1 Criteria for Data Collation

Until now, there is no criterion developed for collating ecotoxicology data necessary in assessing risks of ENMs. Therefore, a guideline developed for evaluating ecotoxicology (ECOTOX) data from the open literature (US EPA 2011), which is specific for pesticides, was used. The guide provides procedures for screening, reviewing, and using both quantitative and qualitative data from the scientific literature on ecotoxicology, and how to apply the information on risk assessment. Here, the procedures are adopted in the context of data collation from relevant scientific publications for use in predictive analysis of behaviour kinetics and toxicity in the emerging domain of nanoecotoxicology.

Guidelines specified in Section 3.2 of ECOTOX (US EPA 2011); meta-analysis notions of aggregating secondary data from various scientific sources (Arnqvist and Wooster 1995); and risk assessment perspectives on physicochemical characterization of ENMs (SCENIHR 2007; Hansen et al. 2007) were applied to undertake the following,

- Develop study criteria for excerpting information from scientific publications.
- Identify input variables from physicochemical properties of $nTiO_2$, properties of water and biological organisms in water that influence behaviour kinetics and toxicity.
- Construct a sample dataset by combining information from different scientific reports into an amenable format for learning.

What type of data should I collect? What are the sources of the data I need? What methods will I use to collate the data? What criteria will I use in data collation? How relevant will be the data collated in meeting the study's objectives? These questions cross the mind of a researcher who is planning to use secondary data in their studies. These questions enabled setting criteria for collating quality data from secondary sources in this study.

Chapters 1 and 2 describe diverse types of ENMs and nanostructures, environmental abiotic and biotic factors that require characterization when conducting exposure and effect tests involving ENMs. Based on this information, a list of pre-selected variables was drawn up to assist in setting criteria for screening and selecting available scientific reports, and extracting relevant data comprised the following,

- Physicochemical properties of ENMs
    - Primary size; size distribution and dispersion media; surface area; shape; purity; crystal phase and composition; zeta potential and dispersive media; source, synthesis method and types of nanoparticles; and surface coating (composition and quantity).
- Environmental abiotic factors
    - pH; ionic strength; concentration of ions, total organic carbon or natural organic matter; temperature; dissolved oxygen; irradiation; total suspended solids; duration of exposure.
- Environmental biotic factors
    - The shortcoming of simulating biotic interactions under laboratory conditions necessitated listing the species, age, size, weight, development stage, feeding, population, and culture of organisms as representative factors.
- Exposure systems used to investigate behaviour kinetics and toxicity of ENMs to indicator aquatic organisms (Table 4.1)

The study aims to utilize secondary data to relate causal factors and responses in an aqueous environment and infer the findings to the aquatic ecosystems. Therefore, the term 'exposure systems' broadly refers to simulated aquatic habitats encompassing water, ENMs, and organisms as environmental processes assigned scores based on the reported descriptors (Table 4.1). Allotting average scores to each descriptor in lieu of desired information establishes the quality of data extracted and applied in this study.

Previous Chapters outline high releases of ENMs into the aquatic environment during use and disposal. For this reason, the criteria score for investigating ENMs found in formulated products compared to pristine forms was high (Table 4.1). Similarly, high scores were allocated to studies that reported using low concentrations of $nTiO_2$ in mixed cultures and flow-through natural water to investigate responses measured under longer durations of exposure (Table 4.1). Acute and chronic tests represent the short and intermediate exposure periods. Rand (1980), cited in Rand et al. (2003), suggested that acute tests should fall within < 10.0% of the organism's lifespan. Accordingly, and in addition to stipulated duration of exposure in standard testing protocols, representative interval criterions were set (Table 4.1).

Table 4.1 Criteria for exposure systems used in assessing scientific reports

Desired information	Broad description of criteria	Score
Exposed ENMs	Pristine and uncoated	1
	Pristine, coated and applicable in nanoproducts	2
	Nanoproducts' formulation	3
Exposure environment (water)	Purified (aqueous)	1
	Standard synthetic	2
	Mesocosms	3
	Natural	4
Exposure methods (analogous to wetlands, lentic and lotic ecosystems)	Static	1
	Semi-static	2
	Flow-through	3
Duration of exposure (days)	≤ 3	1
	3-14	2
	14-42	3
	≥ 42	4
Concentration of ENMs (mg/ ℓ):	≥ 1,000	1
	100-1,000	2
	10-100	3
	1.0-10	4
	0.1-1.0	5
	0.01- 0.1	6
	≤ 0.01	7
Organism community	Axenic (monoculture)	1
	Mixed culture	2

Besides, the predicted environmental concentration of ENMs (refer to Section 2.4.1 in Chapter 2) and the concentration of ENMs determined in environmental samples (refer to Section 1.3) inspired the setting of the highest score listed in Table 4.1. The actual concentration of ENMs in the environment may be lower than the reported values by virtue of the difficulty in distinguishing natural and anthropogenic nanomaterials using metal-basis quantification methods.

4.2.2 Preliminary Assessment of Data

The minimum criteria served as a reference point for assessing published scientific reports with the aim of selecting or excluding them during data collation. Based on the preliminary evaluation of the reports, it was difficult to obtain sources of information that displayed high scores based on the set criteria (Table 4.1). Authors have tested and reported diverse responses using many input variables selected from, (i) different types and nanostructures of ENMs, and (ii) moderately characterized physicochemical properties of ENMs, and environmental abiotic and biotic factors.

Accessible data displays ambiguities in the characterization of ENMs and exposure systems, and description and presentation of findings. The physicochemical properties of ENMs have been characterized using diverse sample preparation methods, testing media and analytical techniques. In some instances, the ENMs were characterized using purified aqueous water while target responses were discussed based on observations in natural or standard synthetic water. Hence, associating properties of ENMs with reported behaviour kinetics and toxicity responses was impractical.

In many instances, partially characterized exposure media stressed specific factors, for instance, either the pH or ionic strength or both, when evaluating target responses. Some reports presented findings based on uncharacterized natural water, such as tap water, which made interpretations difficult. Diverse testing protocols used produced fragmented scientific data describing various aspects of exposure systems, behaviour kinetics, and toxicity of ENMs. Researchers adopted and applied diverse existing protocols, with or without modification, or developed testing procedures in laboratories that limited comparisons. Thus, reports demonstrate numerous responses having diverse units of measurement that reduce the depth of available and extractable information. Moreover, unrealistic concentrations of ENMs, as high as 20,000 mg/ℓ have been reported (Heinlaan et al. 2008).

Designed experiments apply specific objectives to generate primary data published in scientific reports. However, different objectives motivated the use of secondary data in this study. The details of information desired were missing from available sources because of objective-oriented experimental designs, data preprocessing, and partial or inadequate reporting. Moreover, concurrent measurement of behaviour kinetics and toxicity during experiments was unattainable and not reported. However, the problems exhibited by nanoecotoxicology data are not unique. Hitherto, heterogeneous study findings have limited

the statistical analyses on secondary data combined from various scientific reports (Arnqvist and Wooster 1995). In this regard, the following are measures adopted in this study,

- Selection of one compound as case study engineered nanomaterial. Considerable variations in the types and nanostructures of ENMs and measured responses do not lend themselves to comparisons and establishment of meaningful trends.
- Adoption of a flexible approach guided by minimum criteria that allows an expanded list of potential input variables and other relevant information.
- Selection of fewer scientific reports that meet at least 50.0% of the criteria and conduct an in-depth inspection of data to understand preparation techniques and derivation methods used to generate them.
- Collation of data from scientific reports that provided detailed information about ENMs used, exposure conditions and procedures applied.

The aim of adopting this approach is to structure a database representing a broad range of variables drawn from physical, chemical, and biological properties of exposure systems. Additional data sources, such as product catalogues and test protocols, permitted access to, and computation of, unreported information about properties of ENMs and exposure media (especially standard synthetic and purified water). However, it was difficult to approximate properties of partially characterized natural water and ENMs in media unrelated to that used to test reported endpoints. Similarly, mathematical functions and known scientific principles used to generate statistically averaged values in reports aided estimation of continuous data. Thus, based on descriptions of ENMs in previous Chapters, and challenges explained in this Section, data collated were specific to $nTiO_2$ as a representative nanomaterial.

4.2.3 Data Collation and Categorization

Relevant information was collated using external desktop research technique. Using suitable keywords in Google search engine enabled access to data from multiple sources including journal articles, book Chapters, industrial brochures, institutional manuals, and company websites. References in published reports and books provided additional information resources. Data collated manually from texts, tables, supplementary materials and graphs, were keyed in MS Excel worksheet in an amenable format for subsequent analyses. The heterogeneous data were aggregated based on procedures described by Han et al. (2012).

Collated data were quantitative and qualitative. Studies logically described qualitative data that were difficult to measure. Gathering both data types complements and enriches the understanding of multivariate databases (Han et al. 2012), which provides meaning to describing behaviour kinetics and toxicity of $nTiO_2$ in aquatic environment. To strengthen the quantitative data paradigm, data collation integrated descriptive concepts, contexture and interpretation frameworks applied in social and information sciences (Babbie and Mouton 1998; Kaplan and Dduchon 1988; March and Smith 1995; Trochim and Linton 1986).

Conceptualizing theories in inductive learning before subjective information can explain logic from quantitative facts (Falkenhainer and Michalski, 1986). Contextualization in qualitative research assists the discovery of quantitative rules where sample data are extrapolated to define and generalize similar occurrence at larger scales (Han et al., 1993). Interpretation of all relevant information allows their application in explaining phenomena. These frameworks guided the approach of collating both data types. Hereafter, these approaches helped to (i) distinguish dependent and independent variables, and (ii) construct relevant variable lists for analyses and prediction.

Obtaining both data types from texts, tables, and other supplementary materials, was straight forward, but extracting quantitative data from graphs required intense concentration. Requests for raw data used to create graphs from designated authors were not successful. Previously, data from graphs has been extracted by photographically enlarging graphs in published reports and measuring data values using an analogue scientific ruler with an accuracy of 99.0% (Gibbon 1977). Other methods described by Church (2001) include, first, projecting graphs on graph paper wall using overhead transparency and reading off values. Second, photocopying enlarged graphs and using scale transforms of lengths to extract and enter data values in a spreadsheet. Third, scanning graphs and saving them in formats that are readable by applicable software to extract data values.

In this study, the second method suggested by Church (2001) was implemented through enlarging 2D-graphs in scientific reports and printing them on A-4 size paper. The X- and Y-axes of the graphs were calibrated by measuring their lengths and corresponding scales using a mathematical ruler. From a reference point, any data point on the graph was determined using calibrated ratios and keyed in an MS Office spreadsheet. The contrast between data extracted from graphs and similar data points reported in texts and tables of some published reports permitted estimation of the accuracy. Computed relative absolute errors were ≤ 0.035 (or ≥ 96.5% accuracy), suggesting a safe estimate of data that may be

above or below the statistically averaged values presented in the graphs. The method used to excerpt data contributes to the computed errors, but the assumption is the data points still lay within the confidence intervals of raw data used to create the graphs.

Six broad categories define the data collated (Appendix 4.6), comprising first, historical information related to ENMs and experimental designs. Second, static physicochemical properties of $nTiO_2$. Third and fourth, the abiotic and biotic factors, respectively. Fifth, the dynamic properties of $nTiO_2$ in an aqueous environment that represent the behaviour. Sixth, the toxicity effects of $nTiO_2$ on aquatic organisms (Appendix 4.2). The criteria (Table 4.1) and their revision outlined in Section 4.2.1, guided collation of relevant information presented in subsequent Sections.

4.3 Historical Information

The historical category rationally holds detailed information about the source of data and ENMs; sample preparation, exposure methods, and analytical techniques used, and test protocols followed during experiments for reference purpose. The information links collated data to the source and that held in subsequent categories. The listed data assisted in counter-checking and imputing relevant missing data values, deriving additional input variables.

4.4 Static Physicochemical Properties of $nTiO_2$

The word ‘static’ judiciously represents inherent properties of $nTiO_2$ nanostructures as supplied, tested and used in experimental studies. Prevailing conditions in aqueous environment may instigate alteration of these properties without necessarily changing the intrinsic crystal attributes.

4.4.1 Crystal Structure

Titanium dioxide (TiO_2) is a mineral whose crystal structure constitutes elemental compositions with mono-, poly- or amorphous crystallography demonstrating varying properties from different atom arrangement in the lattice (Dakota Matrix Minerals 2012). The bulk form of TiO_2 exhibits amorphous, anatase, brookite, rutile, and ilmenite crystal phases (Table 4.2). The scientific literature classifies nano-forms of TiO_2 using similar crystal phases. Accordingly, this Section describes overlapping properties with distinctions between bulk and nano-forms made where applicable. The structural configurations and thermal stability of the different crystal phases influence molecular bonding and reactivity.

Table 4.2 Crystal phase, system and structural configuration of $nTiO_2$

Crystal phase	Crystal system	Structural configuration	References
Anatase	Tetragonal	Dipyramidal	(Reyes-Coronado et al. 2008; US EPA 2010a; Macwan et al. 2011; Ralph 2012)
Rutile	Tetragonal	Prismatic	(Reyes-Coronado et al. 2008; US EPA 2010a; Macwan et al. 2011; Ralph 2012)
Brookite	Orthorhombic	Dipyramidal	(Reyes-Coronado et al. 2008; US EPA 2010a; Macwan et al. 2011; Ralph 2012)
Amorphous	Mixed	Transitional	(Reyes-Coronado et al. 2008; US EPA 2010a; Macwan et al. 2011; Ralph 2012)
Ilmenite	Trigonal	Rhombohedral	(Reyes-Coronado et al. 2008; US EPA 2010a; Ralph 2012)

Depending on synthesis conditions, the amorphous phase crystallizes to either anatase or rutile (Chen and Mao 2007), anatase stabilizes and transforms to rutile between 550-1,000℃ (Hanaor and Sorrell 2011) and brookite transforms to rutile at > 750℃ (Di Paola et al. 2008). While brookite is less studied and has minimal commercial value, ilmenite is found in ores and poses synthesis challenges for scientific investigation (Raghavender et al. 2013). For these reasons, this study excludes brookite and ilmenite crystal phases. Notably, considerable scientific information and related application of anatase and rutile favour their selection as case study $nTiO_2$ crystal phases in this study. The two materials have found wide application in engineering (for instance, textile, aerospace, and energy conversion), healthcare, sanitation, consumer products, and research among others.

The TiO_2 is a semiconductor that absorbs photons in the visible and ultraviolet (UV) solar spectrum wavelengths. The absorption induces electron excitation that traps titanium cations (Ti^{3+}) and oxygen anions (O^-) in localized sites including grain boundaries (Berger et al. 2005). The instability created by trapped electrons and resultant charge migration enhance reactivity, a feature attributed to photocatalytic application of TiO_2, mainly remediation of organic pollutants and pathogens. The photocatalytic processes involve oxidation of water to hydroxyl free radical ($^{\cdot}OH^-$) and $H^{+,}$ and reduce oxygen to form superoxide radical anions (O_2^-) by the excited electrons. The (O_2^-) and H^+ react with the metal oxide's electrons to form H_2O_2 that is highly reactive and can penetrate organism cells (Banerjee et al., 2006). Thus,

in an aquatic environment, continuous exposure of $nTiO_2$ to UV irradiation portends reactive interactions with beneficial fauna and flora with unidentified consequences.

Anatase and rutile have distinct physicochemical (Ohtani et al. 2010) that explain differences in structure, photon energy absorption, and applications (Di Paola et al. 2008). The wide energy band gap and quantum states govern energy adsorption and photocatalytic properties of TiO_2 crystal phases (Chen and Mao, 2007; Faure et al. 2013). In the pure state, anatase has a larger band-gap (3.2-3.3 eV) compared with rutile's (3.0-3.06 eV)(Li Puma et al. 2008; Reyes-Coronado et al. 2008) that translate to UV-A 376-387 nm and visible light 405-413 nm spectrum regions, respectively. The anatase shows deeper charge carrier excitation path in the crystal lattice, which explains high reactivity compared with rutile (Luttrell et al. 2014). Because of this, anatase is suitable in photocatalytic applications while rutile is a pigment in many commercial products for UV protection.

At the nanoscale, structural configurations of anatase and rutile exhibit differences in locations of reactive atoms and bonding mechanisms. The size and surface free energy determine the thermal stability of $nTiO_2$ (Perego et al. 2009; Diebold 2003). Rutile is the most abundant and thermodynamically stable phase found in ores (US EPA 2010a), and anatase is the most stable phase at the nanoscale (< 4.9 nm) (Di Paola et al. 2008). Accordingly, anatase nanostructures are the main products generated using bottom-up synthesis methods (refer Section 2.1.1) because of reduced average surface energy (Diebold 2003).

At room temperature, stability of anatase and rutile is high in the < 11 nm and > 35 nm size range (Perego et al. 2009), and similarly at low temperatures in crystals < 14 nm and > 30 nm (Zhang and Banfield 1998), respectively. Accordingly, the stable size range suggests prospective coexistence of anatase and rutile in $nTiO_2$ above 14 nm, the incidence of either phase in aqueous environment owing to the size and temperature, and dominance of anatase at low temperatures and vice versa. Thermal stability is a desirable attribute that protects structural deformation of nanostructures in commercial products (Diebold 2003), which also indicates likely fate, transport, and persistence after their release into the environment.

Production and characterization methods, which encompass thermal phase transitions and stability at different temperatures, control the type and composition of crystal phases in $nTiO_2$. For instance, P25 supplied by Evonik Degussa Inc. is a commercial photocatalytic $nTiO_2$ powder widely used in research as a representative nanostructure. Per the

manufacturer, the P25 is a mixture of anatase (80.0%) and rutile (20.0%) produced by flame hydrolysis method. Repeat characterization of P25 shows variations in anatase (80.8 ± 4.8%), rutile (15.3 ± 1.4) and amorphous (3.5 ± 5.9%) compositions in the same batch (Ohtani et al., 2010). Majority scientific reports state the crystal phase compositions as per manufacturer's specifications. Few have reported anatase vs. rutile as 82:18 (Keller et al. 2010), 79:21 (Warheit et al. 2007; Planchon et al. 2013) and 86:14 (Simon-Deckers et al., 2009). Three-phase mixtures of anatase, for example (76.6%), rutile (18.4%) and amorphous (9.0%) (Hartmann et al. 2010), are scarcely reported.

The dominant phase defines the registered crystal structure of $nTiO_2$. For example, anatase coexists in predominantly rutile phase-T-Lite^{TM} SF (Auffan et al. 2010) and rutile in primarily anatase phase-Hombikat UV 100 (Hartmann et al. 2010). The T-Lite^{TM} SF is a highly stable (Wiench et al. 2009) pristine $nTiO_2$ manufactured by BASF used as inorganic UV blocker ingredient in cosmetic products. The Hombikat UV 100 is photocatalytic pristine $nTiO_2$ manufactured by Sachtleben Chemie GmbH. Ohtani et al., (2010) described independent, rather than associative, surface reactivity and photocatalytic properties of anatase and rutile in P25 mixtures. Thus, there are likely disparate effects of the two crystal phases in the event of release into the aquatic environment.

Anatase and rutile crystal phases negatively correlate in coexisting mixtures, which hinders their concurrent use as input variables. However, anatase demonstrates high reactivity (Luttrell et al. 2014) and toxicity to organisms (Ji et al. 2011; Clément et al. 2013; Campos et al. 2013; Mansfield et al. 2015) compared with rutile. Accordingly, this study selects anatase as an input variable representing the crystal structure of $nTiO_2$.

4.4.2 Size, Shape and Surface Area

Production methods introduced in previous Chapters influence the size and shape of $nTiO_2$. Reports demonstrate polydisperse primary and agglomerate sizes, with non-uniform indices as high as 0.6-1.0 (Hartmann et al. 2010; Campos et al. 2013). Considering the P25 as an example once more, the manufacturer characterizes the mixture as 21 nm spherical particles reported as supplied primary size. Independent accounts corroborate the shape but contrast the size of P25 owing to sample preparation and analytical methods applied. Anatase and rutile are described as spherical or ellipsoid crystals (Sadiq et al. 2011) or spherical and cuboid crystals, with nominal sizes of 55 nm and 48 nm, respectively (Planchon et al. 2013). Independently tested, the primary size of P25 reported lies in the range between 18 and 30 nm, inclusive (Li and Sun 2011; Griffitt et al. 2008; Planchon et al. 2013).

The crystal structure of $nTiO_2$ and production methods influences the primary size. For instance, the crystal and primary sizes were comparable for P25 but dissimilar for Hombikat UV 100 because of the amorphous phase found in the structure (Jensen et al. 2004). Besides, mixtures of anatase and rutile synthesized using chlorine process acid hydrolysis yielded spherical (9 nm) and rod-shaped (37 nm) crystals, respectively (Planchon et al. 2013). Therefore, no single size can describe ENMs effectively, and the shape of $nTiO_2$ is dependent on the dominant crystal phase. Thus, it is likely that the differences in sizes, shapes, and composition of crystal phases govern the overall shape of P25 reported by the manufacturer.

Studies commonly report average or median particle sizes of $nTiO_2$, but aqueous dispersions demonstrate polydisperse particle number, volume, or intensity size distributions. Particle number defines density counts of nanoparticles. Volume distribution is size dependent where large sizes have higher distribution and vice versa. Intensity distribution estimates size from optical scatters or diffraction that changes according to the shape of particles. In urban surface water runoff, density number count of $nTiO_2$ < 100 nm was found to be 3.5×10^7 particles/ℓ, while that between 100 and 300 nm was 3.15×10^8 particles/ℓ (Kaegi et al. 2008). In engineered systems, non-destructive image analyses confirm $nTiO_2$ size distributions of 4-30 nm (Westerhoff et al. 2011) and ≤ 50 to ≤ 700 nm (Kiser et al. 2009) in effluent released into freshwater. These findings demonstrate likely retention of intrinsic properties of $nTiO_2$ discharged in freshwater that necessitates selecting a representative variable related to size.

The top-down production methods (refer to Section 2.1.1) can yield ENMs having uneven sizes and structural defects at grain boundaries, which account for the heterogeneous surface characteristics and reactivity (Cao 2004; Hornyak et al. 2009). The irregular crystals formed comprise broken bonds that alter structural configurations of the ENMs. Therefore, the size and shape of ENMs describe the surface-to-volume ratios, reveal packing density of surface atoms and electrons in crystal lattices, and account for surface defects (Grassian 2008; Hornyak et al. 2009; US EPA 2010b).

Size and therefore surface area enhance surface energies and reactivity (Oberdörster 2004). Energy absorption by $nTiO_2$ increases with decreasing the size and vice versa, and decreases with increasing concentration, with peak backscattering observed at 62 nm (Popov et al. 2005), a stable region for rutile crystals often applied as UV blockers. Thus, besides the crystal structure described in the previous Section, the size, and shape of ENMs govern photon charge transfer and reactivity in ENMs.

The size of ENMs is a ubiquitous property not described by single point values. In an aqueous environment, the distribution of ENMs as particle numbers, agglomerates or fractals can shift single point values. Besides, inconsistent units of measuring the sizes render this variable less feasible for modelling. Conversely, surface area, often characterized as specific surface area (SSA), assimilates salient features that ably combine morphology and polydispersity of ENMs in an aqueous environment. Thus, SSA is a relevant variable that represents the size of ENMs, which encompasses partitioning and behaviour kinetics (such as adsorption, catalysis, and reactivity) of diverse nanostructures.

4.4.3 Surface Chemistry

The surface chemistry of ENMs, similar to other charged solids, is the driving force behind their behaviour and reactivity in the environment. Again, production methods influence the surface charges of ENMs (Cao 2004; Hornyak et al. 2009). Owing to their sizes, high energy is required to rebuild broken bonds on ENMs surfaces, leading to metastability (Cao 2004). The surface charge density describes the energy distribution on ENMs that are important in molecular interactions. Poisson's differential equation (Eqn. 4.1) relates surface charge distribution to electrostatic potential, the latter approximated to surface potential using the Poisson-Boltzmann equation (Eqn. 4.2).

$$\nabla^2\psi = -\frac{\sigma}{\varepsilon} \tag{4.1}$$

$$\psi = \psi_0 exp(-\kappa a) \tag{4.2}$$

where ψ is the electric potential, σ is the surface charge density, ε is the permittivity of the media, ψ_0 is the surface potential, a is the diameter of the particle, κ is the Debye-Hückel parameter that associates surface potential to colloid interactions. Thus, smaller particle sizes would depict high surface potential (Eqn. 4.2). The zeta potential is a relative index of surface potential that measures the overall surface energy of charged solids that indicates stability in an aqueous environment (Fig. 4.2).

The Figure illustrates the surrounding of charged nanoparticles in suspension with oppositely charged counter ions strongly bound in the stern layer or weakly bound in the diffuse (weakly layer that stretches into the bulk fluid. The arbitrary lines mark the regions where the surface potential, stern potential, and zeta potential are measured. The interaction between charged nanoparticles and other charged solids create an electrostatic potential on the surface of that particle that is a measured of zeta potential.

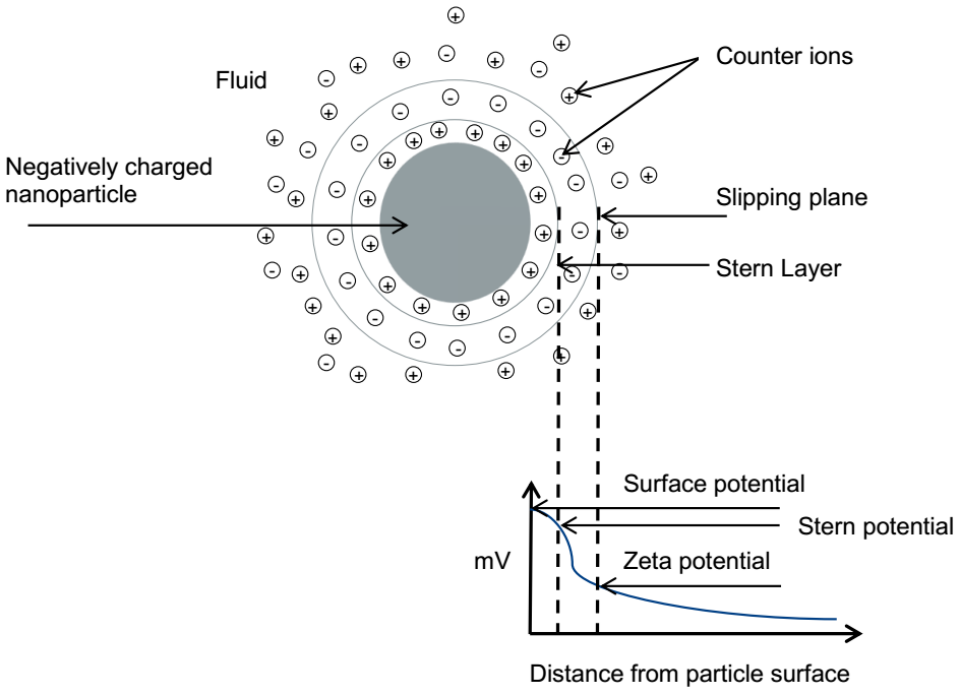


Figure 4.2 An illustration of zeta potential of a charged nanoparticle.

Adapted from: Technical Note, Malvern Instruments (2011).

The Henry's equation relates kinetic mobility of charged solids to the magnitude of zeta potential and properties of the exposure media (Eqn. 4.3).

$$U_e = \frac{2\varepsilon \mathrm{Z} f(\kappa a)}{3\eta} \tag{4.3}$$

where U_e is the electrophoretic mobility, Z is the zeta potential, $f(\kappa a)$ the Henry's function which is a correction factor that relies on the geometry of the particle. The value of $f(\kappa a)$ is 1.5 in an aqueous environment (Smoluchowski's model) and 1.0 in non-aqueous environment (Huckel's model), and η is dynamic viscosity. The Smoluchowski and the Huckel's models are applicable in strong electrolyte dispersions (> 1 mM) of colloids > 200 nm and dilute electrolyte dispersions of colloids < 200 nm, respectively.

Nanostructures of TiO_2 are capped with protective shells to modify their surfaces and improve their application in nanoproducts (such as in cosmetics) (Siddiquey et al., 2007; Jacobs et al., 2010; Lewicka et al., 2011). Capping techniques, such as passivation, introduce barriers on ENMs' surfaces that enclose dangling bonds and protect the movement of unattached electrons. Besides, the modified surfaces shift the point of zero charge (PZC) of ENMs in aqueous suspensions (Huynh and Chen, 2011), enhance the

electrostatic and steric stabilization effects (Cao 2004; Hornyak et al. 2009) and reduce the potential reactivity.

The $nTiO_2$ is capped solely inorganic or combined inorganic and organic agents for specific applications. Examples of inorganic capping agents are, but not limited to, aluminium oxide (Al_2O_3) and aluminium hydroxide ($Al(OH)_3$). In combination, organic coats, such as hydrated silica ($SiO_2.nH_2O$), polydimethylsiloxane (PDMS) $(C_2H_2OSi)_n$, dimethoxydiphenylsilane ($C_{14}H_{16}O_2Si$) and triethoxycaprylylsilane ($C_{14}H_{32}O_3Si$), are grafted onto inorganic shells. The inorganic shells protect nanoparticles from photo-induced reactivity while organic coats improve dispersion of ENMs in formulations, such as cosmetic products.

The scientific literature does not provide adequate quantitative data on capping agents applied. The wide use of pristine uncoated P25 in research partly explains the insufficiency in information. A study reported a high composition of $Al(OH)_3$ (8.5-13.5 w/w%) compared to PDMS (4-6 w/w%) capped on 14-16 nm core T-Lite^{TM} SF (Auffan et al. 2010). The T-Lite^{TM} SF is a mixture of anatase spheres (9.1%), rutile rods (90.9%) (Labille et al. 2010). Considering the irregular sizes and morphology of ENMs discussed in previous sections, quantifying surface coats given their percentages only maybe challenging.

Under favourable conditions, degradation of PDMS exposes $Al(OH)_3$ to reactive processes that alter the structure and surface energy of T-Lite^{TM} SF (Labille et al. 2010; Auffan et al. 2010; Virkutyte et al. 2012). Nevertheless, the end products of PDMS weathering, namely CO_2, SiO_2 and H_2O, are not environmental hazards (Griessbach and Lehmann, 1999). Conversely, likely dissolution, speciation and alteration of $Al(OH)_3$, and other inorganic coats by leaching of Al, may influence reactivity and pose risks in the aquatic environment. The solubility of Al is lowest at pH 6-7 (Gensemer and Playle 1999), increases with increasing pH, and presence of complex ligands at $pH < 6$ and > 8 (Parent et al. 1996).

Dispersion of organic colloids from weathered pristine (Labille et al. 2010) and formulated (Botta et al. 2011) $nTiO_2$ can affect the turbidity of an aqueous environment. Moreover, the reactivity of weathered colloidal matter with $Al(OH)_3$ generates silico-aluminate (Si-O-Al) compounds (Auffan et al. 2010). The compounds are environmentally resistant and thermally stable rigid inorganic geopolymers (Davidovits 1991) that persist in the environment.

Substances designated as process contaminants may double up as stabilizing agents for ENMs. For instance, Al_2O_3 (0.3%) and SiO_2 (0.2%) are impurities on P25 $nTiO_2$ (Product information catalogue, Evonik Degussa Inc.) as well as capping agents on ultrafine $nTiO_2$

with concentrations of 7.0% and 1.0%, respectively (Warheit et al. 2007). This dual negative-positive activity of impurities requires investigating the likely effects on behaviour kinetics and toxicity in the aqueous environment.

Inadequate quantitative information on surface energy of ENMs necessitates selecting surface coating as a categorical variable. Notwithstanding capping materials, surface energies of ENMs vary under different environmental conditions. As a plausible variable, surface coating considerably defines the surface chemistry and reactivity levels of core $nTiO_2$. The diverse characteristics of capping agents affect the net surface charge of metal oxides, hence, the magnitude of zeta potential.

4.5 Environmental Abiotic Properties

4.5.1 pH

The pH is an ion concentration scale that defines colloidal surface charge (Hunter 1981). Surface charges at the liquid-solid or solid-solid interfaces develop through hydrolyzation of metal oxides and de-/protonation of adsorbed organic substrates (Hunter 1981; Elimelech et al. 1995). Ionization of hydrogen and hydroxyl ions (Eqns. 4.4 and 4.5) (Zhao et al., 1993) and adsorption of cationic, anionic and neutral molecules exert control on the net surface charge (Hunter 1981) (Eqn. 4.6).

$$(OH^-) \leftrightarrow_{OH^-}^{H^+} (H_2O) \leftrightarrow_{OH^-}^{H^+} (H_3O^+) \quad \dots \text{ adsorption of water} \tag{4.4}$$

$$(MO^-) \leftrightarrow_{OH^-}^{H^+} (MOH) \leftrightarrow_{OH^-}^{H^+} \left(MOH_2{}^+\right) \dots \text{ metal surface hydroxyl} \tag{4.5}$$

$$\sigma_0 = F(\Gamma_{H+} - \Gamma_{OH-}) \tag{4.6}$$

where MO is metal oxide, σ_0 is the net surface charge, and Γ_{H+} and Γ_{H-} are adsorption densities. More protons are lost when the pH increases resulting in increased negative charge (refer to Eqn. 4.6). Hydrolyzation of $nTiO_2$ depends on pH variations. In aqueous environment, bulk TiO_2 is an intermediate oxide with a net zero charge at pH 5.6-6.2 (Elimelech et al., 1995), interchangeably referred to as the PZC or isoelectric point (IEP). Potentiometric and electrokinetic surface conductivity of charged solids influence the PZC and IEP values, respectively that are equal in the absence of adsorbing ligands. Exclusively, increasing the difference between pH and PZC enhances metal oxides stability (Kissa 1999).

The synthesis methods, size and crystal phase (Chen and Mao, 2007), and particle concentration (Suttiponparnit et al. 2010) influence the PZC of $nTiO_2$. The PZC of 3.6-8.1 nm

$nTiO_2$ was found to be 4.6-6.5 (Finnegan et al. 2007), which falls within the PZC region of bulk-TiO_2. In purified aqueous water, the increase in pH was higher in the presence of smaller particles (5 nm) compared with larger ones (32 nm) (Pettibone et al., 2008). Using similar exposure media, a 33-fold increase in $nTiO_2$ concentration decreased the pH by 10.50% (Suttiponparnit et al. 2010). Additionally, the crystallinity of $nTiO_2$ nanostructures controls pH change (Finnegan et al. 2007; Hartmann et al. 2010). The Nernst's model (Eqn. 4.7) demonstrates the influence of pH on surface potential around the PZC of rutile (Larson and Attard, 2000).

$$\psi_0 = \frac{k_B T}{ze} In \frac{[pH]}{[pH]_{PZC}} \tag{4.7}$$

where ψ_0 is the surface potential, k_B the Boltzmann constant, T the absolute temperature, and ze the proton charge. The pH controls solubility and reactivity of organic and inorganic compounds, which influence the surface chemistry of charged solids by changing the surface charge density. Oxidation and reduction in acidic and basic conditions, increase the positive and negative surface charges, respectively that stabilize $nTiO_2$ in aqueous environment (Martyanov et al. 2003). Often, scientific reports provided static information on pH conditions but seldom during tests. Unreported pH values were averaged based on information found in test protocols applied.

4.5.2 Ionic Strength

Dissolved ions conduct electricity in the aqueous environment. Electrical conductance is an indirect measure of the ionic strength or concentration of electric charges. The ions that determine ionic strength comprise Li^+, Na^+, NH_4^+, K^+, Mg^{2+}, Mn^{2+}, Ca^{2+}, Cu^{2+}, Fe^{2+}, Zn^{2+}, Al^{3+}, Fe^{3+}, Cl^-, SO_4^{2-}, HCO_3^-, CO_3^{2-}, F^-, NO_3, HSO_4^-, OH^-, $NaSO_4^-$, and $NaSO_3^-$ (Mccleskey et al., 2012). In sub-surface water, the Lind (Eqn. 4.8) and Russell (Eqn. 4.9) models estimate ionic strength from measured conductance.

$$I_c = 1.4769 \times 10^{-5} K_s + 0.00015 \tag{4.8}$$

$$I_c = 1.5 \times 10^{-5} K_s \tag{4.9}$$

where I_c is the ionic strength, and K_s the corrected specific conductance measured at 25℃. Under ideal conditions that disregard interactions of ions, ionic strength is equivalent to the actual concentration of dissolved ions (Eqn. 4.10).

$$I_c = 0.5 \sum_{i=1}^{k} c_i z_i^2 \tag{4.10}$$

where c_i is the molar concentration of specific ions, and z_i the ions' charge determined by the valence. Under definite conditions, ions interact and their effective concentration is lower than the actual one. The activity coefficients define interaction of ions that correct their electrostatic capability. The ions' charge significantly affects the magnitude of ionic strength as well as the activity. The Debye-Hückel parameter (Eqn. 4.11) associates ionic strength due to multiple dissolved salts to zeta potential. This parameter demonstrates the influence of ions on surface charge and can estimate their activity coefficients.

$$\kappa = \sqrt{\left(\frac{e^2}{\epsilon k T} \times \sum_{i=1}^{k} c_i z_i^2 \right)} \tag{4.11}$$

where e^2 is the electron charge. Under ideal conditions, ionic strength influences the stability of charged solids by controlling the zeta potential. Conductance, hence, electrostatic forces of interacting phases, induces either electro kinetic cohesion or adhesion that transfers negative or positive charge effects. In dilute electrolytes, negatively charged oxides, for example, nTiO_2, acquire positive and negative charges below and above the PZC, respectively, which decreases attractive forces in monovalent regimes. The magnitude of surface charge decreases towards the PZC as the strength of electrolytes increases.

In view of the various test protocols used, a broad range of types and concentration of inorganic matter were potential predictors and their application would increase the models' dimensionality. The measures adopted to reduce dimensionality were; first, classify inorganic matter in terms of ion speciation and valence effects on the surface energy of ENMs. In an aquatic environment, micro- and macro elements would have combined rather than independent effects on ENMs unlike in designed laboratory experiments. Thus, types of inorganic matter signify ionic constituents grouped as a single qualitative input variable, designated as 'charged ions' to differentiate exposure systems.

Second, the concentration of inorganic matter partially relates to ionic strength. In the absence of reported ionic strength values, the *eq.* was estimated using stoichiometric concentration of ions and elements found in added salts using PHREEQC® (Parkhust and Appelo 1999) at reported pH and temperature. In addition, the ionic strength was computed by substituting reported values of conductance in the Lind's equation (Eqn. 4.8).

4.5.3 Organic Matter

In freshwater, organic matter are complex polydisperse and amorphous ligands from natural and anthropogenic sources. Aquatic ligands originate from allogenic (ex-situ) and autogenic (in-situ) sources consisting unique physical, chemical, and biological components. Factors that influence properties of aquatic ligands include their origin, environmental conditions (such as pH, ionic strength, temperature, and solar radiation), photoactivity and microbial activity and adsorption. Solubility, molecular size and weight, composition, structure and surface chemistry differentiate properties of organic matter.

In natural water, concepts classify organic and inorganic matter into dissolved solutes (≤ 1 nm) and colloids (≤ 450 nm), and particulates (≥ 450-100,000 nm) (Buffle and Leppard 1995; Rand et al. 2003). Properties of filtered (usually (≤ 450 nm or less) and unfiltered dissolved colloids differ and have a have a bearing on measured reactivity. Recently, organic matter has been categorized as dissolved matter (≤ 1 nm), colloids (1-1,000 nm) and particles (≥ 1,000 nm)(Aiken et al., 2011). The classification has been revised based on advances in characterizing primary structures of dissolved molecules, mainly aquatic humus substances, that exhibit low molecular weights between (200-2,000 Daltons (Da)) (Leenheer et al. 2001; Sleighter and Hatcher 2007).

Inorganic matter, fulvic materials, and biopolymers are types of colloids that play a role in fate and transport of aquatic pollutants (Buffle et al. 1998). Structures of inorganic matter, for instance, amorphous metal oxides, kaolinites, and illites, have a broad range of PZC's (pH 2.3-9.4) and surface areas (10-800 m^2/g) while organic matter comprises fulvic, and rigid and flexible biopolymer substances (Buffle et al. 1998). Fulvic substances form a large part of natural organic matter (NOM) that have high molecular mass (500-5,000 Da) (Beckett et al. 1987; Chin et al. 1994) and high surface charge densities (-6 to 11 *meq.*/g) of dissolved sites (Buffle et al. 1998). Molar mass of rigid biopolymers is very high (10^4 to $>10^5$ Da) with surface densities of -0.35 to -6 *meq.*/g, whereas that for flexible biopolymers is low (500-800 Da) with charge densities of -2.5 to -5.5 *meq.*/g (Buffle et al. 1998).

Under favourable aquatic conditions, organic molecules dissociate forming anionic charge densities that control surface energy, behaviour, and the fate of interacting solids (such as ENMs). Molecular properties of organic matter and water chemistry buttress the strength and magnitude of surface charges. In surface water and sediments, dissolved molecules constitute humus (50.0%), low molar carboxylic and phenolic acids (≤ 25.0%), neutral

compounds (≤ 15.0%), and other species (≤ 10.0%) (Vanloon and Duffy 2005). Negatively charged humus comprises fulvic acid (FA), humic acid (HA) and humin substances, which define their equilibrium charges and alkalinity in water (Eby 2004; Vanloon and Duffy 2005).

Total organic carbon (TOC) in FA and HA ranges from 40.7 to 58.7 $(w/w)\%$ (Eby 2004), with average content in the acids estimated as 50.6 and 57.2 $(w/w)\%$[2], respectively. It is problematic to characterize the complex nature of NOM. Concentration of TOC, including carbonic acid, represent organic matter. In subsurface water, NOM is 1.7 times the concentration of TOC, a conversion factor determined based on the highest possible content (58.7 $w/w\%$) of carbon (Vanloon and Duffy 2005).

The subsurface TOC and NOM relationship conversion factor was modified (Eqn. 4.12) based on TOC values in simulated aquatic conditions reported in the literature. The composition of TOC in surrogate organic matter (that is Suwanee River FA, HA, and NOM) was 52.0-55.0 $(w/w)\%$. The denominator value in Eqn. 4.12 is the sum of the highest possible content, and the minimum and maximum data of TOC in reported studies. The assumption is that this higher value would accommodate related data uncertainties.

$$Organic\ matter = \left[100\left(\frac{1}{58.7}+\frac{1}{52.0}+\frac{1}{55.0}\right)\div 3\right] * TOC \approx 1.8 * TOC \qquad (4.12)$$

Background TOC in academic grade purified water is 0.005-0.01 mg/ℓ (Material data sheet, Millipore) and ≤ 20 ppb (Product certificate of analysis, Fisher Scientific). Applying the conversion unit (Eqn. 4.12) estimates *eq*. NOM in deionized (0.009-0.018 mg/ℓ) and Optima (0.036 mg/ℓ) water solvents. Total carbon includes inorganic carbon content (CO_2/H_2CO_3, HCO_3^- and CO_3^{2-} species) from added salts in standard synthetic media and background organic carbon in purified water from biological and metabolic sources.

Organic matter affects turbidity of water, hence, the colour, taste, odour, air exchange, as well as changes in temperature, irradiation, and pH. In the presence of ENMs, a 22-fold increase in the concentration of oxalic acid was found to lower the pH by 37.5% (Pettibone et al., 2008). Moreover, organic matter influences the behaviour kinetics and migration of contaminants and subsequent interaction with organisms. In toxicity tests, carbon as a source of nutrients to organisms requires investigation for associated effects to ENMs.

[2] http://www.humicsubstances.org/elements.html

Feasible variables are types of organic matter, which embody distinctive properties of ligands and concentration of TOC (or *eq.* NOM) in each volume of exposure media. The two aspects are determinants of behaviour kinetics and toxicity of ENMs investigated.

4.5.4 Temperature

Temperature plays a significant role in physical and biochemical reactions that can affect behaviour kinetics and toxicity of ENMs in an aqueous environment. Notable examples of physicochemical properties affected by temperature include the density, pH, salinity, and conductivity; dissociation and partitioning of chemical compounds; dissolution of oxygen; metabolism, adsorption, redox- and biological reactions.

An increase in temperature increases dissociation rates and reduces oxygen solubility (Mihelcic and Zimmerman 2010). Moreover, physiological reactions, metabolism, and detoxification in aquatic organisms depend on temperature variations (Nikinmaa 2014). The permittivity of pure water decreases with decreasing temperature, which indirectly affects the zeta potential and in turn stability of charged solids. The Arrhenius equation relates chemical reaction rates to the absolute temperature (Eqn. 4.13).

$$k_r = A_f e^{-(E_a/RT)} \tag{4.13}$$

wherek_r is the rate constant for respective reaction, A_f is the pre-exponential factor related to collisions and steric factors with similar units to k_r, E_a is the activation energy, and R is the gas constant. The aquatic habitats sustain specific species of fauna and flora based on their tolerance to existing temperatures (Cairns et al. 1975). The aquatic organisms thrive best in conditions simulated in laboratory cultures to promote their metabolism, for example, algae (16-27°C) and *D. magna* (18-22°C). Conversely, the location, habitats and their depths, seasonal variations, solar radiation, and turbidity influence temperature magnitudes.

4.5.5 Concentration of $nTiO_2$

Based on their size, ENMs can be classified as dissolved substances, colloids, and particles of inorganic origin. Because of this, the concentration of $nTiO_2$ falls outside intrinsic properties suitable to be an abiotic factor. Concentration defines the ratio of the mass of a substance per unit volume of aqueous media. Scientific reports described concentration in terms of mass, particle number, and cluster of nanostructures. Owing to inadequate information on dynamics of particle numbers in an aqueous environment, information

collated was inclined towards the mass aspect of concentration. Moreover, metrics of mass concentration explain particle kinetic principles in an aqueous environment.

Experiments conducted in batch reactors applied mass concentration-related principles to explain behaviour kinetics and toxicity of $nTiO_2$. Thus, it is important to mention dynamic mechanisms related to pollutant transport in the aqueous environment. Advection and diffusion densities are mechanisms that stratify mass fluxes linked to discharge sources, transfer quantities, and transport kinetics. The advection mass flux density estimates the migration of a substance from a point over a given area per given time based on the flow of water (Eqn. 4.14). Conversely, the diffusion mass flux density is a function of concentration gradients described by Fick's Law (Eqn. 4.15), which explains the mass transfer of molecules across boundaries towards achieving an equilibrium state.

$$J = C.v \tag{4.14}$$

$$J = -D\frac{dC}{dx} \tag{4.15}$$

where J is the mass flux density, C the concentration, v is the velocity of water in the direction of flow, D the diffusion coefficient, and $\frac{dC}{dx}$ in the concentration gradient in the environment. The random motion of substances reduces concentration differences of molecules in water. Fick's Law partially explained the transport $nTiO_2$ in porous media but failed to distinguish the concentration gradients of charged and neutral molecules (Chowdhury et al. 2011). Equation 4.16 estimates total mass flux of compounds in fluids, such as water.

$$\dot{m} = J.A \tag{4.16}$$

where $\dot{m}$ the total mass flux across a boundary, and A is the area of the boundary. The aquatic environment considered determines steady state pollutant concentrations achieved. It is probable to attain equilibrium concentrations of chemical substances in batch reactors. However, pseudo equilibrium concentrations are likely in aquatic environment, such as rivers. In rivers, advection flux density supersedes diffusion flux density in estimating mass flows of chemicals downstream and accumulation in littoral and even subsurface zones.

A study reported low solubility (< 0.02-1 nM) of filtered (≤ 200 nm sieve) $nTiO_2$ (8.5 ± 2 to 37 ± 5 nm) in aqueous environment at pH < 5.2 to 8.2, ionic strength (5-8 mM) and *eq.* organic matter (0.2-5 mg/ℓ) (Planchon et al. 2013). Thus, ionic $nTiO_2$ would be negligible to find in

aquatic environment except for insoluble colloids and particles. Section 4.2.1 presents concentrations of $nTiO_2$ analyzed in environmental samples while Section 4.5.1 describes the influence of concentration on the pH.

Knowledge about mass loads of ENMs is relevant in risk assessment but limited in explaining bioavailability and toxicity effects based on the crystal structure, size, morphology, and surface chemistry. It is imperative to link mass loads of ENMs with their size-related interfacial interactions and toxicity. In purified aqueous media, a high concentration of smaller nanoparticles displayed high surface area per unit mass concentration (Suttiponparnit et al. 2010), which expresses the notion of surface area concentration (SAC). The SAC was found to be a valuable metric in assessing pulmonary toxicology of insoluble ENMs (Fissan et al. 2006) and its rationale in environmental and biological samples has recently been reported (Palchoudhury et al. 2015).

Specific surface area (SSA) of ENMs (refer to Section 4.4.2) is characterized using specific mass doses and adsorption methods, for example, BET (Table 2.1). Application of mass-based concentration of ENMs suggests variations in SSA. The ECHA (2012) recommends derivation of volumetric-based SAC using SSA to assess the toxicity of ENMs. Because of this, the derived SAC (m^2/ℓ or m^{-1}) of $nTiO_2$ was a product of reported values of SSA (m^2/g) and concentration (mg/ℓ). As an integrated physical property, the SAC reduces dimensionality and difficulties associated with polydispersity, particle numbers, and morphology of ENMs. Practical application of SAC in ecotoxicology would be to evaluate selected endpoints based on surface area as opposed to ubiquitous sizes of ENMs. Reported invariable intrinsic surface area and reactivity of $nTiO_2$ after clustering (Sun and Lee 2012) supports the suitability of SAC as a surface area-to-volume variable with the m^2/ℓ units adapted in this study.

4.5.6 Duration of Exposure

Duration of exposure estimates periods within which specific endpoints that elucidate behaviour of ENMs, and direct and indirect effects to biological organisms are measured. Moreover, this variable determines the potential time taken to respond in the event of incidental and accidental releases of ENMs in aquatic environment. Mass loads and residence time of chemical compounds in aquatic environments' boundaries affect concentration fluxes, partitioning, accumulation, bioavailability, uptake and toxicity of ENMs. Acute and chronic exposures represent residence time for assessing behaviour, fate, and

toxicity of ENMs. The first-order kinetics explains concentration decay in batch reactors (Eqn. 4.17).

$$C_t = C_o e^{-kt} \tag{4.17}$$

where C_t is the concentration at a given point in time (t) after release of chemical pollutants, C_o the influent concentration, k the decay rate, and t is the residence time in each system. The ENMs can be removed in aquatic environment because of treatment methods, gravity induced settlement, uptake and accumulation in biological organisms or build-up in aquatic compartments, among others. Kinetic decay functions describe pollutant removal processes.

4.6 Behaviour Kinetics and Toxicity of ENMs in Water

This Section presents information collated on the interdependent influence of static physicochemical properties and abiotic factors on interactions of $nTiO_2$ and organic matter to adsorption and aggregation behaviour of ENMs. Also presented are relevant underlying principles governing adsorption and aggregation of charged solids, adopted to explain similar aspects in nanoecotoxicology. Sections 4.6.3 and 4.6.4 describe biotic factors and behaviour kinetics of ENMs influencing toxicity effects. Tables 4.3 and 4.4 and Appendix 4.2-4.5 summarize relevant information compiled from collated data on properties of ENMs, exposure conditions, behaviour kinetics, and toxicity of ENMs for use in this study.

4.6.1 Adsorption of Organic Matter on $nTiO_2$

As an introduction, the words adsorptive, adsorbent and adsorbate have the following meanings; a substance in aqueous solution, an adsorbing solid surface and the mass or molecules attached on the solid, respectively. Theory and application of adsorption can be found in Dąbrowski (2001). After their release into an aquatic environment, ENMs would coexist with other macro-, micro-, and nano-scale natural, biological and anthropogenic matter that determine resultant behaviour kinetics, and in turn toxicity. Thus, ENMs may act as adsorbents or adsorptive with other ligands in aqueous dispersion. This study considers $nTiO_2$ as the adsorbent and various types of organic matter as the adsorptive.

Sections 4.4 and 4.5 discuss factors that influence adsorption behaviour of ENMs in an aqueous environment. Physical (Van der Waals attraction), chemical (covalent and metallic) and ionic (electrostatic) forces govern bonding mechanisms at interfacial sites. A review by Philippe and Schaumann (2014) on adsorption mechanisms controlling aqueous phase adsorption of organic matter on natural and engineered nanomaterials reveals knowledge

gaps in current understanding fate and transport of ENMs in natural water. Nanoecotoxicology studies have used organic surrogates to estimate the possible behaviour of ENMs in an aquatic environment but likely complex adsorption mechanisms are unknown.

The existence of multiple competing adsorptive in aquatic systems implies intricate adsorption mechanisms with ENMs. The surface charges of ENMs may be non-uniform and heterogeneous, a characteristic of charged solid surfaces complexed by adsorbates from diverse ligands in natural water (Jaroniec et al. 1991; Elimelech et al. 1995). Moreover, the presence of inorganic ions would affect surface charge densities leading to stability or a shift in the point of zero charge (Hunter 1981), thus, potential reactivity and partitioning.

Organic matter is one of the competing adsorptive in the aquatic environment. Properties of organic adsorptive (refer to Section 4.5.3) would govern their surface interaction with ENMs and equilibrium transfer of adsorbates. Favourable conditions promote dissociation of organic molecules and formation of functional groups that participate in surface reactions. In $nTiO_2$-organic matter interactions, the ENM gains protons from deprotonated functional groups. For instance, deprotonation of carboxyl $(-COOH)$ and phenol $(-OH)$ forms carboxylate $(-COO^-)$ and phenolate $(-O^-)$ anions, respectively (Eqn. 4.18 and 4.19). The bonding of anions (or conjugate bases) with ENMs alters net surface charge of the latter.

$$-COOH \leftrightarrow -COO^- + H^+ \ldots.\text{carboxyl deprotonation} \quad (4.18)$$

$$-OH \leftrightarrow -O^- + H^+ \ldots.\text{ phenol deprotonation} \quad (4.19)$$

The adsorbates on $nTiO_2$ were found to be carboxyl and phenol groups of humic acid (Yang et al. 2009; Erhayem and Sohn 2014b). Moreover, irreversible adsorption was attributed to chemical bonding of carboxylate (Sun and Lee 2012) and citrate (Mudunkotuwa and Grassian, 2010) on $nTiO_2$. Boltzmann equation describes the dissociation of functional groups in an aqueous environment (Eqn. 4.20).

$$K_a = \frac{v_-}{v_0}[H^+]exp\left(\frac{-e\psi_0}{k_B T}\right) \quad (4.20)$$

where K_a is the dissociation constant of functional groups adsorbed on the surface of charged solids, $v_-$and v_0 are the number of negative and neutral functional groups at the interaction sites having different proton concentration, respectively; e is the electron charge; ψ_0 is the surface potential, k_B is the Boltzmann constant, and T is the temperature. The magnitude of temperature influences thermal energy between interacting surfaces.

The Langmuir (Araujo et al. 2005; Pettibone et al. 2008), Freundlich (Erhayem and Sohn 2014b; Chekli et al. 2015), Brunauer-Emmet-Teller (Mudunkotuwa and Grassian 2010; Barreto et al. 2012), and Langmuir-Freundlich (Sun and Lee 2012) isotherm models have been used to explain adsorption of organic molecules on $nTiO_2$. A review by Yang and Xing (2010) provides useful insights on bonding mechanisms and isotherm modelling of aqueous phase adsorption of organic matter on ENMs, with a focus on carbon nanotubes (refer to Section 1.1). The potential to use diverse thermodynamic models to explain adsorption underscore the influential role of properties of ENMs, adsorptive and aqueous environment discussed in Sections 4.4 and 4.5.

In summary, the size and crystal phase of $nTiO_2$ (Pettibone et al. 2008), pH (Chen and Elimelech 2008; Yang et al. 2009; Mudunkotuwa and Grassian 2010; Erhayem and Sohn 2014b) and ionic strength (Erhayem and Sohn 2014b; Chekli et al. 2015; Chekli et al. 2015) affect adsorption rates and capacity. The surface area and atomic configuration of $nTiO_2$ crystals avail adsorption sites for adsorbate binding. The pH and ionic strength regulate the surface chemistry of ENMs, dissociation, and transfer of adsorptive from the aqueous environment to the adsorbing surface.

The valence, concentration and binding mechanisms of mono- and polyatomic ions determine the quality of adsorbates formed through co-adsorption and bridging (Murphy et al. 1994; Sun and Lee 2012; Chowdhury et al. 2012; Erhayem and Sohn 2014a; Erhayem and Sohn 2014b; Chekli et al. 2015). The initial concentration of organic matter (Sun and Lee 2012; Chekli et al. 2015; Chekli et al. 2015) and types of organic matter (Pettibone et al. 2008; Erhayem and Sohn 2014a; Erhayem and Sohn 2014b) determine adsorption mechanisms and partitioning on adsorbents. In the aquatic environment, an inverse relationship exists between partitioning coefficients of organic adsorptive and concentration of solid adsorbents (Site 2001).

Authors (Pettibone et al. 2008; Mudunkotuwa and Grassian 2010) reported using 2-(N-morpholino) ethanesulfonic (MES) and 4-2(-hydroxyethyl)-1-piperazineethanesulfonic (HEPES) pH control buffers in their experiments. There are indications that buffers influence aggregation of nanoparticles (Stemig et al. 2014) and compete with humic acids on $nTiO_2$ (Erhayem and Sohn 2014a). Thus, will incorporate buffers as components of evaluated surrogate organic matter. Table 4.3 and Appendix 4.2 details specific information excerpted from six published scientific reports on adsorption of organic matter on $nTiO_2$ particles relevant for this study.

Table 4.3 Summary of data on adsorption of organic matter on $nTiO_2$ in water

Influencing variables	Description of adsorption findings	Reference
5 nm and 32 nm anatase, and 30 nm rutile particles. pH 2.0, 5.5 (MES) and 6.5 (HEPES) in purified aqueous media.	Size influences adsorption rates but not adsorption capacity for oxalic and adipic acids irrespective of the pH. No increased adsorbate affinity observed based on size. At all pH conditions, the comparable adsorption capacity of oxalic acid on 5 and 32 nm particles that is different for adipic acid. Differences in structural configurations, reactivity and high surface bonding of anatase with oxalate, explain reduced and increased adsorbates on 30 nm rutile compared to 32 nm anatase, respectively.	(Pettibone et al. 2008)
4 nm anatase, pH 2.0, 6.0 (MES) and 7.5 (HEPES) in purified aqueous media.	Citrate adsorbates decreased with increasing pH owing to enhanced repulsion between negatively charged $nTiO_2$ and polyatomic citrate anions in basic conditions.	(Mudunkotuwa and Grassian 2010)
P25 (17.7 ± 6.1 nm). pH 5.0 and 7.0, 10 mM in purified aqueous media.	At pH 5 and Ca^{2+} conditions, adsorbates were 2.5-fold more than in same pH and K^+ regime and 1.1-fold in pH 7 and Ca^{2+} conditions.	(Chowdhury et al. 2012)
P25 (ca. 27 ± 4 nm). pH 4.8 and 7.8, 10-100 mM, Na^+, purified water. pH 7.9 and 4.6 mM, MHW. pH 4.8-10.0, 10, 20, 40 mM, K^+, Na^+, Ca^{2+}, Mg^{2+}, Cl^-, NO_3^-, $H_2PO_4^-$, and HCO_3^- in purified aqueous media.	Under all test conditions, adsorption of SRHA was > SRNOM > SRFA. Adsorbate concentration increased and reduced in mono- and di-valent regimes, respectively. Adsorption constants for all types of organic matter decreased with increasing pH and increased with increasing ionic strength in purified water. Adsorption constants are higher in MHW than in purified water owing to the presence of polyatomic counterions that adsorb on the ENMs surface; enhance repulsion and decrease adsorption with increasing ionic strength. Antagonism between polyatomic anions of functional groups and the highly negative surface of $nTiO_2$ increases the repulsive energy that reduces adsorbate quantities in basic conditions.	(Erhayem and Sohn 2014b)
P25 (21 nm). pH 6.53-8.04, 1.96-735 mM, DOC (2.65-10.01 mg/ℓ), Ca^{2+} (0.06-10.30 mM) in river, lake, ground, and seawater media.	Independently and in a combination of the influencing factors, adsorbed amount decreased with decreasing initial concentration of DOC, increasing pH, Ca^{2+} content and ionic strength, and vice versa. The significance of ionic strength to adsorbed mass manifested after combining with initial concentrations of DOC and Ca^{2+}.	(Chekli et al. 2015; Chekli et al. 2015)

4.6.2 Aggregation of $nTiO_2$ in Water

In published literature, agglomeration and aggregation are words that interchangeably describe clustering of charged solids. In the context of this study, agglomeration and aggregation designate particle-particle and particle-other solid clustering in an aqueous environment, respectively. In fluids, Brownian collisions, attachment, disintegration, and regrowth define the disorderly manner by which charged solids approach and aggregate. The Smoluchowski's population model theoretically explains the collision frequency of approaching particles as the product of second-order aggregation reaction rate and concentration of respective particles (Eqn. 4.21).

$$J_{ij} = K_{ij} n_i n_j \tag{4.21}$$

where J_{ij} is the collision frequency between the *i*-th and *j*-th particles, K_{ij} is the second-order collision rate constant, and n_i and n_j is the concentration of *i*-th and *j*-th particles, respectively. However, the continuous collisions characterized by the Smoluchowski model do not address hydrodynamic and colloid interaction factors that affect timescales of colloid diffusion in fluids. Therefore, the model is strengthened by including (i) perikinetic aggregation rate, K_{ij}^{p}, a Brownian motion-diffusion rate from Stokes-Einstein model (Eqn. 4.22), (ii) orthokinetic aggregation rate, K_{ij}^{o} from Camp and Stein model that considers fixed particle concentration in laminar drag flows (Eqn. 4.23), and (iii) differential sedimentation rate, K_{ij}^{ds}, a Stokes' Law model explaining settlement of particles and clusters (Eqn. 4.24).

$$K_{ij}^{p} = \frac{2k_B T}{3\mu} \frac{\left(a_i + a_j\right)^2}{a_i a_j} \tag{4.22}$$

$$K_{ij}^{o} = \frac{4}{3} G\left(a_i + a_j\right)^3 \tag{4.23}$$

$$K_{ij}^{ds} = \left(\frac{2\pi g}{9\mu}\right)(p_s - \rho)\left(a_i + a_j\right)^3\left(a_i - a_j\right) \tag{4.24}$$

where k_B is the Boltzmann constant, T is the temperature, μ is the fluid viscosity, a_i and a_j are the radii of *i*-th and *j*-th particles, respectively, G is the shear rate, g is the gravitational acceleration, and p_s and ρ is the density of particles and fluid, respectively. The classical and extended Smoluchowski models help to understand the theoretical aggregation kinetics of charged solids adopted in nanoecotoxicology (Domingos et al. 2009).

In addition, colloid interaction factors involve thermodynamic as well as attractive and repulsive forces between approaching particles. Section 4.6.1 highlights thermodynamic models applied to estimate adsorptive interaction with ENMs and surface transfer of ligands that lead to clustering. The random attachment and growth of solids is driven by diffusion-limited (fast) and reaction-limited (slow) forces that produce permeable and impermeable structures, respectively (Buffle and Leppard 1995). Moreover, the classical and extended Derjaguin-Landau-Verwey-Overbeek (DLVO) (Derjaguin and Landau 1941; Verwey and Overbeek 1948) theory explains the attractive and repulsive forces contributing to aggregation of $nTiO_2$ particles (Liu et al. 2011; Zhou et al. 2013; Zhang et al. 2015).

The theoretical mechanistic models are limited in explaining aggregation of ENMs because of size dispersity, and heterogeneous surface charges and morphology from adsorbed ligands (Jaroniec et al. 1991; Elimelech et al. 1995; Bhattacharjee et al. 1998). The adsorbates control the stability of ENMs in both acidic and basic exposure conditions, which makes it challenging to quantify aggregation kinetics with certainty. On this account, factors that influence adsorption (Section 4.6.1) play a significant role in the clustering of ENMs and the aspect of hetero-aggregation is considered. The application of physicochemical properties discussed in Sections 4.4 and 4.5 in a data-driven model would provide a good understanding of likely interactions that influence aggregation in an aqueous environment.

The physicochemical properties that influence aggregation of $nTiO_2$ are primary size (Pettibone et al. 2008; Hartmann et al. 2010; Chowdhury et al. 2013; Seitz et al. 2013) and crystal phase (Liu et al. 2011; Zhou et al. 2013; Iswarya et al. 2015). Additionally, surface coating (Labille et al. 2010; Botta et al. 2011) and concentration (Pettibone et al., 2008; Keller et al. 2010; Sillanpää et al. 2011; Campos et al. 2013; Jacobasch et al. 2014; Mansfield et al. 2015) of $nTiO_2$ influence aggregation.

Particles with the smallest primary sizes exhibit low critical coagulation concentration (CCC) (Zhou et al. 2013; Chowdhury et al. 2013) and their unstable surfaces provide adsorption sites for ligands leading to large aggregate sizes. In addition, large aggregates form with increasing concentrations of $nTiO_2$ due to high collisions and reduced surface energy. Besides, the dominant crystal phase determines the surface charge of $nTiO_2$ (Lebrette et al. 2004; French et al. 2009; Hartmann et al. 2010) and subsequent surface interactions. Studies on citrate-stabilized gold nanoparticles demonstrated enhanced aggregation in the presence of organic matter attributable to the capping agent (Nason et al. 2012).

The abiotic factors that govern aggregation rates and quality of fractals include pH, ionic strength, concentration of organic matter, charged ions and duration of exposure (Domingos et al. 2009; Keller et al. 2010; Li and Sun 2011; Ottofuelling et al. 2011; Chowdhury et al. 2012; Chowdhury et al. 2013; Zhu et al. 2014; Chekli et al., 2015; Chekli et al., 2015). Aggregation decreases with decreasing pH and increasing ionic strength (Pettibone et al. 2008; Chowdhury et al. 2013). Whereas low concentration of organic matter promotes aggregation, high concentration sterically stabilizes $nTiO_2$ (Labille et al. 2010).

An increase in pH and concentration of organic matter lowers the electrophoretic mobility (EPM) of ENMs and enhances steric stabilization, respectively, due to repulsive forces between hydroxyl and adsorbed functional groups. Moreover, increasing ionic strength compresses the electric double layer (EDL) of ENMs, which reduces the Debye distance, lowers zeta potential, and promotes aggregation due to electrostatic attraction. Low EPM enhances aggregation of ENMs in natural waters characterized by high ionic strength and low concentration of organic matter (Keller et al. 2010; Ottofuelling et al. 2011). Moreover, a low and high content of organic matter induce diffusion- and reaction-limited aggregation with increasing ionic strength, respectively (Zhu et al. 2014).

Charged ions and their valence can reverse aggregation profiles of ENMs through electrostatic attraction or intermolecular bridging with functional groups. The CCC of divalent ions is lower than monovalent ions. Thus, divalent ions promote aggregation at low pH and ionic strength (French et al. 2009), and also with increasing ionic strength and concentration of organic matter (Zhu et al. 2014) by neutralizing negatively charged adsorbates-$nTiO_2$ that reduce the zeta potential (Zhang et al. 2009). In addition, divalent ions promote fast aggregation by compressing the EDL and reducing the EPM (Chowdhury et al. 2012). At low pH, high concentration of monovalent ions can shift the PZC and neutralize surface charges of ENMs leading to stable nanoparticles (Ottofuelling et al. 2011).

The concentration of organic matter stabilizes ENMs depending on the molecular weight (Nason et al. 2012); thus, the significance of adsorptive properties on aggregation. The bonding mechanisms of short bonds of low molecular weight organic matter adsorbed on $nTiO_2$ promote stability and aggregation of ENMs in low and high ionic strength conditions, respectively (Labille et al. 2010). In contrast, a decrease in intramolecular charge repulsion of large molecular weight organic matter enhance aggregation (Labille et al. 2010). Table 4.4 and Appendix 4.3 provide details of information collated from thirteen published reports and are applicable in this study.

Table 4.4 Summary of data on aggregation of $nTiO_2$ in water

Influencing variables	Description of aggregation findings	Reference
P25 (uncoated 27 ± 4 nm mixture), seawater (pH 707 mM, TOC (0.65 mg/ℓ), mesocosm freshwater (pH 7.18 mM, TOC, 63.45 mg/ℓ)	High ionic strength in low NOM seawater triggers rapid aggregation with increasing concentration of $nTiO_2$. Low ionic strength in high NOM mesocosm freshwater enhances stability at all concentrations.	(Keller et al. 2010)
Hombikat UV100 (< 10 nm, anatase and amorphous). P25 (ca. 30 nm, anatase, rutile and amorphous). Hombitan LW-S (ca. 300 nm, anatase) in algal test media.	Particles < 10 nm form large aggregate sizes compared to those formed by particles *ca.* 30 and 300 nm irrespective of the duration of exposure and crystal phase. Ions in exposure media and particle ratios enhance aggregation at low concentrations of ENMs.	(Hartmann et al. 2010)
P25 (19.8 nm). Details of natural and standard synthetic water used as exposure media listed in Appendix 4.3.	Stability influenced by magnitude zeta potential, pH, ionic strength (IS) and concentration of DOC in aqueous media. Low IS, high DOC, and high zeta potential in peat bog water reduced aggregation and enhanced stability. High IS, low DOC, and low zeta potential in seawater promoted aggregation and reduced stability. Low DOC content and moderately high IS influenced low stability in EPA hard water.	(Ottofuelling et al. 2011)
P25 (21 nm). Brackish (104 and 145 mM) and lake water (0.43 and 0.45 mM). Details listed in Appendix 4.3.	At low particle concentration (1.00 mg/ℓ), organic matter content and adsorption interactions irrespective of ionic strength influenced aggregation size in brackish (≤ 500 nm) and lake water (≤ 200 nm). A 100-fold increase in concentration promotes aggregation in brackish water (≤ 3,000 nm) and lake water (≤ 1,500 nm), enhanced in former due to high ionic strength.	(Sillanpää et al. 2011)
P25 (18.4 ± 6.0 nm), 1.00 and 10.00 mM (KCl) in purified aqueous media.	At pH 5, a 10-fold increase in ionic strength promoted high aggregation. At pH 7, aggregation increased irrespective of ionic strength and $nTiO_2$ concentration.	(Chowdhury et al. 2011).
P25 (18.4 ± 6.0 nm), pH 4, 6 and 8, fulvic acid (FA) (0.00-5.00 mg/ℓ), 5.00 mM (NaCl) and Fe^{3+} (0.0-0.2 mM) in purified aqueous media.	At all pH ranges, FA reduced aggregation. Trivalent Fe^{3+} and FA stabilized $nTiO_2$ in acidic conditions but promoted aggregation at *ca.* PZC and basic conditions due to electrostatic bridging.	(Li and Sun, 2011)

Table 4.4 Summary of data on aggregation of $nTiO_2$ in water (*continued*)

Influencing variables	Description of aggregation findings	Reference
P25 (17.7 ± 6.1 nm). pH 5.0 and 7.0, 10 mM (KCl, $CaCl_2$), and humic acid (HA) (1.00 mg/ℓ) in purified aqueous media.	In the absence of HA, aggregation in Ca^{2+} regime is higher than in K^+ regimes at both pH conditions. At pH 5, adsorbed organic matter induces high surface charge density that reduces aggregation and aggregate sizes in Ca^{2+} regimes. Reduced charge density in K^+ regimes has similar effects. In the presence of organic matter, Ca^{2+} reduces negative surface charge while K^+ increases the charge by a factor of 3, which reduces stability.	(Chowdhury et al. 2012)
P25 (25.1±8.2 nm), $nTiO_2$ (0.1, 0.25, 0.5, 200, 800 and 1,500 mg/ℓ) in standard moderate hard water.	Increasing concentration of ENMs increases aggregation. Stable aggregate sizes in concentrations ≤ 0.5 mg/ℓ and large sizes in concentrations ≥ 200 mg /ℓ observed.	(Ma et al. 2012)
NM-103 and NM-104 (ca. 20 nm coated rutile) and P25 (21 nm) in standard hard water.	A 10-fold increase in the concentration of $nTiO_2$ increases aggregation size by 120.0-322.0% irrespective of surface coating and crystal phases. Aggregation size of P25 was 1.6 and 1.3-fold, and 1.3 and 1.9-fold higher than coated NM-103 and NM-104 at 1.0 and 10 mg/ℓ concentrations, respectively. Polydispersity indices of heterogeneous aggregates formed were 0.36-0.89 and 0.45-0.91 at 1.0 and 10 mg/ℓ concentrations, respectively.	(Campos et al. 2013)
P25 (21 nm), A100 (6 nm), DOC (8.1 mg/ℓ) in standard hard water.	Aggregate sizes formed in semi-static systems were 1.2-2.3-fold those formed in flow-through systems. Brownian motion of ENMs in the latter system promotes fast aggregation and disaggregation of unstable structures.	(Seitz et al. 2013)
$nTiO_2$ particles (< 40 nm mixture, 1.0 and 10 mg/ℓ) in standard hard water.	In static exposure systems, a 10-fold increase in the concentration of ENMs increased aggregate sizes by 2.5-fold.	(Kim et al. 2014)
P25 (21 nm), pH 6.53-8.04, 1.96-735 mM, DOC (2.65-10.01 mg/ℓ), Ca^{2+} (0.06-10.30 mM) in river, lake, ground, and seawater.	High ionic strength and low organic matter content promote fast aggregation that forms large sized unstable aggregate structures of loosely packed fractals. Low ionic strength and high organic matter content promote slow aggregation that forms stable aggregate structures owing to adsorption. High-density aggregate fractals formed.	(Chekli et al. 2015; Chekli et al. 2015)

4.6.3 Toxicity of $nTiO_2$ on Algae

In the aquatic environment, algae are found to be sensitive to toxicants and widely used in dose-response risk assessments for representative ecosystems (Wright and Welbourn

2009). Algae are autotrophs that produce energy through photosynthetic reactions. Community interactions of biological organisms influence their exposure to chemical pollutants and in turn the toxicity. Hence, effects of toxicants to algae would be evident in higher organisms in the aquatic food web owing to trophic transfer.

In most cases, effects of $nTiO_2$ to algae are assessed using homogenous axenic or monocultures of algae species, with or without similar tests conducted using other species or different taxa- such as fish. A comparable axenic study established that toxicity of $nTiO_2$ was contingent on algae species (Cardinale et al. 2012). In natural water, diverse organisms may influence toxicity through additive or antagonizing effects (Mihelcic and Zimmerman, 2010), which are inadequately tested endpoints for ENMs.

A study using mesocosm freshwater demonstrated potential neutralization of $nTiO_2$ toxicity to algae species, with enhanced biomass growth attributed to suppressed competition with other organisms (Kulacki and Cardinale 2012). Moreover, effects of $nTiO_2$ on metabolism in algae in mesocosms are specific to exposed species (Cardinale et al. 2012). Mesocosms simulate marginal conditions of aquatic environment that demonstrate mixed cultures. With potential correlation between algae species and heterogeneous cultures of organisms and toxicity of $nTiO_2$, the two aspects are biotic predictors investigated in this study.

The size, crystal phase and increasing concentration of ENMs and duration of exposure were found to inhibit growth rate, thus, biomass growth in algae (Hund-Rinke and Simon 2006; Warheit et al. 2007; Aruoja et al. 2009; Hartmann et al. 2010; Ji et al. 2011; Cardinale et al. 2012; Kulacki and Cardinale 2012; Clément et al. 2013; Iswarya et al. 2015). Photocatalytic effects to algae were shown after lengthy exposure to pre-illuminated $nTiO_2$ (Hund-Rinke and Simon 2006). However, the major link between the photocatalytic generation of reactive oxygen species (ROS) of irradiated $nTiO_2$ and toxicity effects observed in algae species has not been established (Ji et al. 2011; Fu et al. 2015). Marked inhibition of algae growth observed in negative control experiments (Fu et al. 2015) reveals the role of irradiation as a hindrance to the growth rate in the absence of ENMs, which has not been sufficiently tested.

The mechanisms of uptake and toxicokinetics of $nTiO_2$ in algae are not well understood. However, imaging techniques have shown encapsulation of $nTiO_2$ aggregates on algae that inhibited metabolism (Aruoja et al. 2009; Hartmann et al. 2010; Sadiq et al. 2011; Ji et al. 2011; Metzler et al. 2011). Hence, factors contributing to behaviour kinetics of ENMs in

aqueous media would significantly affect toxicity. Appendix 4.4 presents information gathered from five studies on the toxicity of ENMs to algae that are applicable in this study.

4.6.4 Toxicity of $nTiO_2$ on *Daphnia magna*

The water flea *Daphnia magna* is the most sensitive crustacean in freshwater toxicity tests (Wright and Welbourn 2009). Water fleas are primary consumers inhabiting surface and sub-surface regions of the pelagic aquatic environment and are easy to culture in laboratories for toxicity assessments that investigate water quality (Jonczyk and Gilron 2005). The *D. magna* is a species of the genus *Daphnia Sp*. that tolerates hard water, commonly found in lakes and ponds (Ebert 2005; Jonczyk and Gilron 2005). Projections from substance flow analysis (SFA) and multimedia models (refer to Sections 2.4.1 and 2.4.2 in Chapter 2) demonstrated the buildup of ENMs in littoral and pelagic regions that predispose daphnids to suspended and settled toxicants. Moreover, batch reactors used in laboratory studies also simulate ecological habitats for *D. magna*.

The economics of using daphnids is their short lifespans that depend on environmental factors, which permit toxicity tests at different life cycle stages (Jonczyk and Gilron 2005). Under optimum environmental conditions, *D. magna* has four development stages and 6-22 adult instars after hatching; namely, neonates (2-3 days), juvenile (3-6 days), adolescent (6-14 days) and adult (14 to ≤ 90 days) (Jonczyk and Gilron 2005). The sensitivity of organisms to toxicants depend on the development stage, with the egg and adult stages being the least sensitive (Wright and Welbourn 2009). Given the number of instars that complete the life cycle of *D. magna*, relevant biotic predictors that would influence the toxicity of ENMs during exposure are organism age and development stage.

With a life cycle of close to three months, and a large number of adult instars during which the organism reproduces, the uptake, bioaccumulation, and disposition of ENMs can affect physiological processes in offspring. Population collapse in the sixth generation (Jacobasch et al. 2014) demonstrates the significance of long-term effects of $nTiO_2$ to the individual growth and reproductive capability of *D. magna* offspring. As a biotic predictor, organism generation can gauge long-term risk assessment of ENMs in aquatic ecosystems.

Three aspects of feeding in *D. magna* that portend toxicity effects, namely nutrition; depletion; and ingestion, are precursors to detoxification and bioaccumulation of ENMs in biological organisms. In freshwater toxicity tests, nourishment of biological organisms has been found to alter toxicity findings (Rand et al. 2003). The levels of available food sustain

growth and reproduction of zooplanktons (Antunes et al. 2004). Aquatic organisms become vulnerable to toxicity effects when ENMs bind with their food sources. As heterotrophic filter feeders, daphnids feed on algae and other detritus matter (such as bacteria and debris ≤ 70,000 nm) (Ebert 2005). The size of food filtered by the organism falls within the range exhibited by inorganic $nTiO_2$ likely ingested during feeding. Thus, toxicity effects of ENMs on algae and bacteria would manifest in *D. magna*.

The abundance of food levels not only raises the chances of ingesting ENMs (Campos et al. 2013) but also detoxification (Zhu et al. 2010), which reduces bioaccumulation. When food is scarce and binds with ENMs, resultant flocculation and settlement lessens nourishment (Campos et al. 2013) and cleansing capabilities (Zhu et al. 2010), leading to food deficiency and bioaccumulation effects, respectively. Thus, food is an important biotic predictor denoted as 'organism feeding' to represent the underlined aspects influencing reproduction.

The adsorption of ENMs onto organisms is an uptake pathway that was found to hinder molting (Dabrunz et al. 2011), induce behaviour changes (Lovern et al. 2007), and weaken food filtration efficiency (Kim et al. 2010; Zhu et al. 2010; Jacobasch et al. 2014). The weakening of the exoskeleton and inability to ingest food lowers growth leading to mortality. In the aquatic environment, change in behaviour would expose the organism to predation. Another uptake mechanism is ingestion and accumulation of free or bound ENMs in the gut of *D. magna* (Kim et al. 2010; Zhu et al. 2010; Campos et al. 2013; Jacobasch et al. 2014), which can transfer nanoparticles in the food chain.

Toxicity effects manifested in the organism's reproduction depend on size (Warheit et al. 2007; Kim et al. 2010; Dabrunz et al. 2011; Clément et al. 2013; Seitz et al. 2014) and crystal phase (Clément et al. 2013; Campos et al. 2013; Seitz et al. 2013; Seitz et al. 2014) of ENMs. A rise in concentration of ENMs (Warheit et al. 2007; Kim et al. 2010; Zhu et al. 2010; Clément et al. 2013; Kim et al. 2014; Jacobasch et al. 2014) and duration of exposure (Zhu et al. 2010; Kim et al. 2010; Dabrunz et al. 2011; Clément et al. 2013; Kim et al. 2014) increased the inhibition to reproductive capability.

Moreover, the toxicity effects were higher in semi-static exposure systems (Seitz et al. 2013; Kim et al. 2014) than flow-through (Seitz et al. 2013) and static (Kim et al. 2014) systems because of enhanced bioavailability of dispersed and settled ENMs, respectively. Appendix 4.5 illustrates information compiled from seven published reports on toxicity of ENMs to *D. magna*.

4.7 Sample Dataset for Learning

Continuous data summarized in reports were derived using mathematical functions and scientific principles. Relevant but not reported data were manually imputed using values obtained from companies' websites, product catalogues, and published reports on similar ENMs or test protocols. The dataset structure implemented equates similar properties found in purified, standard synthetic, mesocosm, and natural water. Units of measurement were converted and harmonized to attain consistency.

4.7.1 Summary of Predictor and Response Variables

Input variables comprise three physicochemical properties of ENMs, seven abiotic factors, and five biotic factors. The properties of ENMs include composition of anatase crystals (%), representing the structure, configuration, and reactivity, the surface area concentration (SAC) (m^2/ℓ) and surface coating (unitless). The SAC integrates size, morphology, and concentration variables while the surface coating represents the surface chemistry of ENMs.

The environmental abiotic factors comprise charged ions (unitless); duration of exposure (hours); ionic strength (mM); and concentration of organic matter (as natural organic matter (NOM) or total carbon (TC) (mg/ℓ)). The other factors include pH (unitless); temperature (^oC); and types of organic matter (unitless) representing diverse properties of organic matter. The biotic factors include organism age (days in instar life cycle); development stage (neonate and juvenile); feeding (unitless); generation (first to sixth); and culture (axenic or mixed) denoting species diversity in habitats. The organism age and development stage are merged to create a single categorical variable named 'organism growth phase.'

The nature of collated information and units of measuring endpoints determined the selection and derivation of target responses. The mass of organic adsorbates (mg/g of $nTiO_2$) and hydrodynamic size (nm) of $nTiO_2$ are responses selected from adsorption and aggregation endpoints, respectively. The toxicity responses derived are inhibition to (i) biomass growth in algae (based on fluorescence, mass, cell density and % units), and (ii) reproduction in *D. magna* (founded on number of offspring, eggs produced and % units). The biomass growth is an output of growth rates, and inhibition to the later would reduce the expected abundance of algae. Besides, reproduction in *D. magna* determines the population growth rates, hence, inhibition to the former affects the latter.

Applying a formulae by Abbott (1925) or described in OECD (2011) normalized the derived toxicity response values to the negative control values to create a standard unit, namely, percent effect. The negative controls test the health of organisms in experiments and signify baseline conditions for thriving without pollutants, in this case, $nTiO_2$ while normalized values define the effects arising from exposure. The normalized negative cases imply an increase in toxicity effects corresponding to the control and vice versa with normalized positive cases. This study considers values that show an increase in toxicity only.

4.7.2 Statistical Description of the Sample Dataset

A structured sample dataset for learning includes eight quantitative and seven qualitative input variables, four target responses, and 1,036 irregular training examples created by data gaps in collated information. The dataset columns and rows comprise predictor and response matrices, and continuous and nominal values, respectively. The mixed data, amalgamated from 23 published reports, were generated using purified (41.3%), standard synthetic (28.9%), mesocosm (15.9%), and natural (13.9%) water exposure systems. The irregularities necessitated grouping the response variables into four independent subsets (Table 4.5).

Table 4.5 Statistical descriptions of responses' data subsets

Response variables	Training examples	Mean	Median	Minimum	Maximum	Standard deviation
Adsorbed mass (mg/g of $nTiO_2$)	366	13.13 ± 1.42	8.05	0.14	71.42	13.80
Hydrodynamic size (nm)	290	2,659.10 ± 617.85	1,124.25	131.80	27,665.90	5,345.80
Inhibition to algae biomass growth (%)	226	39.16 ± 3.73	33.00	1.64	97.00	28.44
Inhibition to reproduction in *D. magna* (%)	154	58.73 ± 5.82	56.86	1.32	100.00	36.53

The central tendency of continuous data shows distribution around the mean (or median). The variances suggest the diversity of a broad range of training examples measured under different experimental designs (Table 4.5). Thermodynamic models fitted to experimental data may explain the small variance observed in adsorbed mass. The hydrodynamic size exhibits a large variance attributed to different exposure systems and analytical techniques used, with 91.0% of values distributed in the 130-5,000 nm range. The normalized toxicity response values show low dispersion.

Tables 4.6 and Figures 4.3-4.5 illustrate the spread and plots of data values for continuous variables. The dispersion is considerably high for SAC, ionic strength (IS), duration of exposure and concentration of organic matter (NOM), low for anatase and total carbon (TC), and near invariant in temperature and pH variables. Comparatively, the pH, temperature, ionic strength and organic carbon (which refers to the total, dissolved and particulate organic carbon) in natural water depends on the system (Table 4.7 and Appendix 4.3).

Table 4.6 Statistical representation of continuous variables in the sample dataset

Continuous input variables	Mean	Median	Minimum	Maximum	Standard deviation
Anatase (%)	83.70 ± 0.85	82.00	0.00	100.00	13.94
SAC ($m^2/ℓ$)	114.46 ± 15.87	2.00	5.8E-04	876.00	260.26
pH	6.92 ± 0.01	7.60	2.00	8.40	1.60
IS (mM)	72.78 ± 11.61	8.23	0.30	750.00	190.45
Temperature (^oC)	22.43 ± 0.16	22.00	18.00	25.00	2.61
Duration of exposure (hours)	106.55 ± 9.50	36.00	0.00	600.00	155.74
TC (mg/ℓ)	33.53 ± 2.63	13.73	2.15	63.45	26.08
NOM (mg/ℓ)	52.10 ± 9.36	10.01	0.01	917.08	122.07

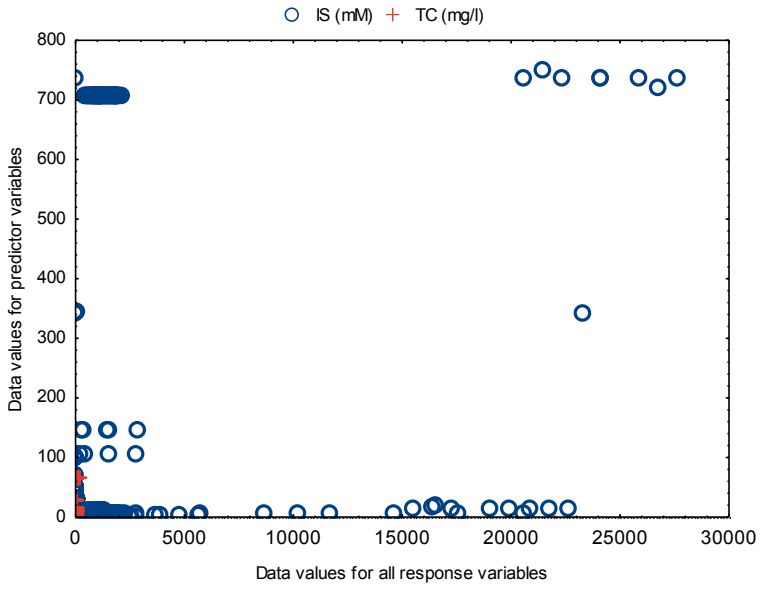


Figure 4.3 Scatter plot of data for all target responses versus ionic strength and total carbon

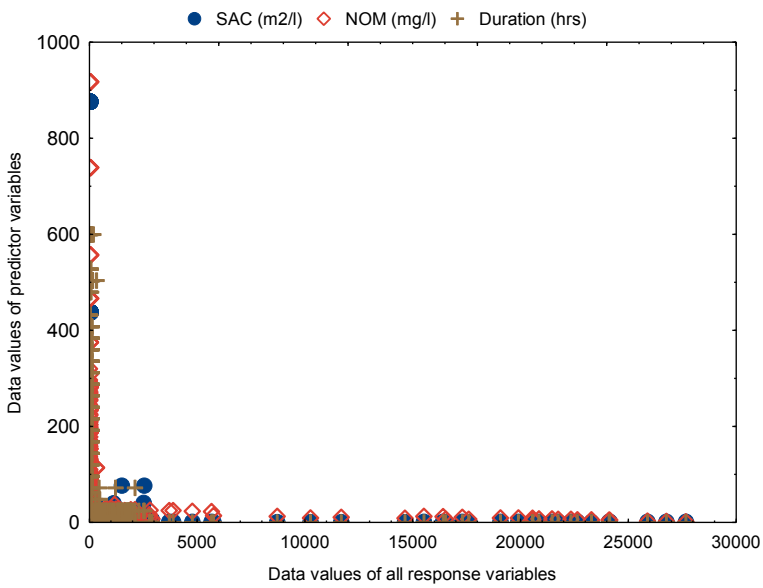


Figure 4.4 Scatter plot of data for all target responses versus SAC, NOM, and duration of exposure.

Except for ionic strength, continuous data displays a narrow range of values. The diverse experimental designs (Appendix 4.2-4.3) contribute to the large variances computed for SAC, IS, duration of exposure and NOM (Table 4.6, and Fig. 4.3 and 4.4). The IS, pH and temperature values in the dataset simulate those established in aquatic conditions (Table 4.6 and 4.7, and Fig. 4.3 and 4.5).

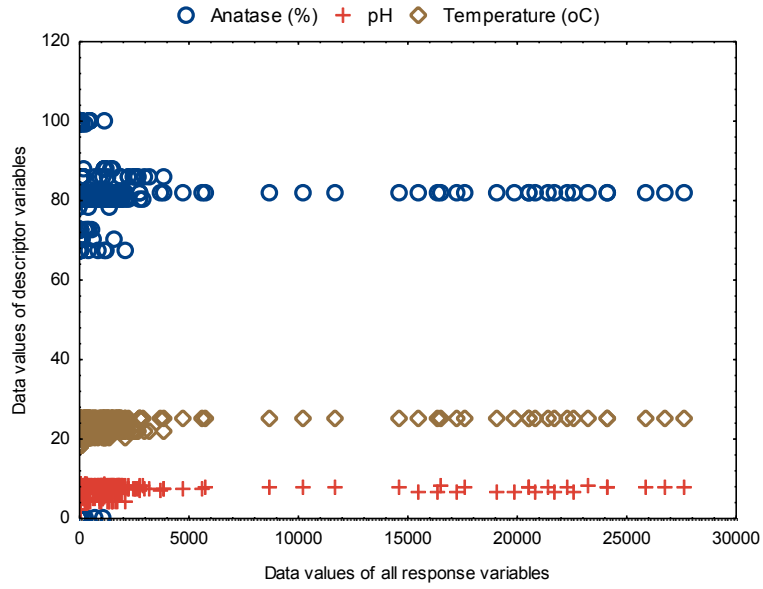


Figure 4.5 Scatter plot of data for all target responses versus anatase, pH, and temperature.

Natural water has a narrow range of pH values, with extreme values caused by human factors (such as the release of industrial pollutants) while temperatures marginally change without extreme external factors. From the dataset, the TC values lie within the expected values in natural water but the maximum value of NOM (~1.00 g/ℓ) exceeds anticipated values (Tables 4.6 and 4.7, and Appendix 4.2-4.5).

Table 4.7 The properties of natural water and analyzed $nTiO_2$ concentration

Variables	Description	Reference
pH	4.00-10.00	(Sala et al. 2000)
Ionic strength (mM)	Surface water: 1.00-5.00 Portable and groundwater: 1.00-2.00 Seawater: 700	aqion[3]
Temperature (^{o}C)	0.0-40.0	(Mihelcic and Zimmerman 2010)
Total organic carbon (TOC) (mg/ℓ) NOM = 1.8 × TOC (Eqn. 4.12)	Seawater: 0.01-1.0 Coastal water: 1.0-3.0 Large rivers: 2.0-10.0 Lake water: 1.0-50.0	(Rand et al. 2003)
ca. Concentration of $nTiO_2$ (mg/ℓ)	Rivers: 0.01-0.011 Urban water run-off: 0.6-1.0 Effluent into surface water: 2.7×10^{-6}-0.042	(Neal et al., 2011) (Kaegi et al. 2008) (Kiser et al. 2009; Westerhoff et al. 2011; Khosravi et al. 2012)

Categorical data are predominantly nominal, except one that has partial interval values. The variables include (i) charged ions in exposure systems, surface coating of ENMs and types of organic matter used in experiments; (ii) species of algae and exposure culture; and (iii) development stage, age, feeding and generation of *D. magna*. The species and culture are dichotomous predictors while the rest are polytomous with 3 to 31 different levels.

Studies reported investigating the influence of ions in exposure media using monovalent K^{+} (2.2%) and Na^{+} (33.0%), divalent Ca^{2+} (1.2%) and Mg^{2+} (0.6%), mono- and trivalent Na^{+} and Fe^{3+} (4.3%), and diverse mixed valent (58.7%) ions. Thirty-one different categories

[3] http://www.aqion.de/site/69

distinguish diverse values of ions and concentrations even in similar aqueous media, which accentuates complexity of charged ions as an input variable. Properties, concentrations and speciation of charged ions, micro- and macro elements in the natural water, microcosm, and various preparations of simulated water portray unique influence on surface energy, behaviour kinetics, and bioavailability of ENMs.

From the dataset, the ENMs used in experiments were uncapped (98.5%) or capped with Al_2O_3 (0.4%), Al_2O_3 and SiO_2 (0.2%), $Al(OH)_3$ and PDMS (0.4%), and $Al(OH)_3$, PDMS, glycerin and other organics (0.5%) (refer to Section 4.4.3). Table 4.8 summarizes 12 different types of organic compounds tested against responses in experimental studies. The SRFA, SRHA, and SRNOM represent surrogate organic matter having unique properties, hence, influence on responses based on their origin. Total carbon and biological organism differentiate aggregates formed during exposure of *D. magna* to ENMs.

Table 4.8 Summary of organic materials in the sample dataset

Types of organic materials in exposure systems	% (out of 100)
Citric acid	1.9
HEPES added to citric acid	1.0
HEPES added to oxalic acid	4.3
MES added to adipic acid	9.9
MES added to citric acid	1.0
Natural organic matter (NOM) (in natural and purified water)	33.4
Oxalic acid	3.2
Suwanee River fulvic acid (SRFA)	8.9
Suwanee River humic acid (SRHA)	4.8
Suwanee River aquatic NOM (SRNOM)	4.3
Total carbon (TC)	23.4
Total carbon and biological organism	4.1

The algae species were cells (58.5%), filamentous (14.4%) and colony forming (10.7%), which represent different ecological habitats in the aquatic environment. Culture, which simulates communities of biological organisms was axenic (57.0%) and mixed (43.0%). Studies reported exposing *D. magna* to $nTiO_2$ as neonates (73.0%) and juveniles (27.0%), two development stages of the organism's 4-5 instar life cycle (refer to Section 4.6.4). The

ages of *D. magna* at exposure were ≤ 24-hours (34.0%), 6 to 24-hours (3.0%), and 72-hours (64.0%). Studies reported investigating toxicity responses using twelve species of algae (Table 4.9) show more use of *Pseudokirchneriella subcapitata*, and one species, *Daphnia magna,* of the genus Daphnia (Tables 4.5-4.6). From the dataset, toxicity was measured in the first (93.7%), second (1.3%), third (1.3%), fourth (1.3%), fifth (1.1%) and sixth (1.3%) generation of *D. magna.* The feeding variable (Table 4.10) represents nutrition to *D. magna* during tests and signifies diverse habitats where type, depletion, or abundance of food would influence the degree of ENMs uptake and subsequent toxicity.

Table 4.9 Summary of information on algae species in the sample dataset

Algae species	% (out of 100)	Algae species	% (out of 100)
Tabularia fasciculata	3.5	*Oscillatoria spp.*	8.9
Anabaena spp.	0.4	*Selenastrum minutum*	4.4
Stigeoclonium tenue	1.8	*Chlorella vulgaris*	11.1
Planothidium lanceolatum	6.2	*Chlamydomonas moewusii*	8.9
Spyrogyra communis	3.5	*Scenedesmus quadricauda*	17.3
Navicula subminiscula	5.8	*Pseudokirchneriella subcapitata*	28.3

Table 4.10 Information about quality and quantity of food for *D. magna* in the sample dataset

Types and quantity of food for *D. magna* during tests	% (out of 100)
0.2 mg C per of *Scenedesmus obliquus* algae per organism per day	18.8
YCT and 1 x 10^7 cells/mℓ of *Selenastrum capricornutum* algae per day	65.6
1 × 10^6 cells/mℓ of *Pseudokirchneriella subcapitata* algae per day	2.6
1.5 mg C/ℓ (*eq.* 5 x 10^5 cells/mℓ of *C. vulgaris* algae)	2.6
0.2 mg C of *Desmodesmus sp*. algae per day	8.4
0.3 mg C/ℓ (*eq.* 1 x 10^5 cells/mℓ of *C. vulgaris* algae)	2.0

4.7.3 Quality of Collated Data

Based on the criteria developed (refer to Table 4.1), data compiled in the sample dataset exhibits a score of 1.71 or 44.7% of the expected maximum quality (Table 4.11). From the database, uncoated and coated pristine nanoparticles mainly used to investigate behaviour kinetics and toxicity endpoints were those not found in formulated products. In respect of the

high ranking assigned to ENMs found in formulated products, the data posted the lowest (33.8%) of the expected quality (Table 4.11).

Table 4.11 Data quality scores based on developed criteria

Desired information	Minimum score	Maximum score	Quality score (% of expected score)
Exposed ENMs	1	3	1.01 (33.8)
Exposure media (water)	1	4	2.02 (50.6)
Exposure methods	1	3	1.49 (49.8)
Duration of exposure	1	4	1.41 (35.2)
Concentration of ENMs (mg/ ℓ)	1	7	2.88 (41.1)
Organism community (representing habitats)	1	2	1.43 (71.3)
Average score of data quality	1	3.83	1.71 (44.7)

The majority of the studies reported investigating behaviour kinetics and toxicity of $nTiO_2$ using simulated (standard or purified) aqueous water (70.2%) as opposed to natural and pseudo-natural mesocosm freshwater (29.8%). With a score of 2.02 of expected quality (Table 4.11), the findings on kinetic properties and toxicity of ENMs have a 50.6% implication to the aquatic environment. Moreover, studies reported using static (57.0%), semi-static (37.0%), and flow-through (6.0%) exposure systems to simulate aquatic conditions, with an average score of 1.49 or 49.8% quality (Table 4.11). The exposure systems were either renewed or not renewed with ENMs or water.

A six-generation toxicity study (Jacobasch et al. 2014) took 126 days to complete but exposure period per generation was 21 days. Because of this, the maximum duration of exposure investigated was 25-days. Duration of exposure accounts for mainly short- and intermediate-term effects of ENMs to biological organisms, and short periods that evaluate adsorption and aggregation behaviour of ENMs. With regard to long-term effects in the aquatic environment, the criteria score is the second lowest at 35.2% of the expected quality (Table 4.11).

A notable criterion is the concentration of $nTiO_2$ investigated against target responses (Fig. 4.6). From the dataset, exposure concentrations (mg/ℓ) used were ≤ 0.01 (0.0%), 0.01-0.1 (1.2%), 0.1-1.0 (6.4%), 1.0-10.0 (25.8%), 10.0-100.0 (34.4%), 100.0-1,000.0 (10.5%) and

1,000.0-4,000.0 (21.8%). The concentrations of nanoparticles used in experimental studies, which affect the magnitude of SAC (refer to Section 4.5.5), are higher than those analyzed in environmental samples with only 7.6% ≤ 1.0 mg/ℓ (refer to Section 1.3 and Table 4.7).

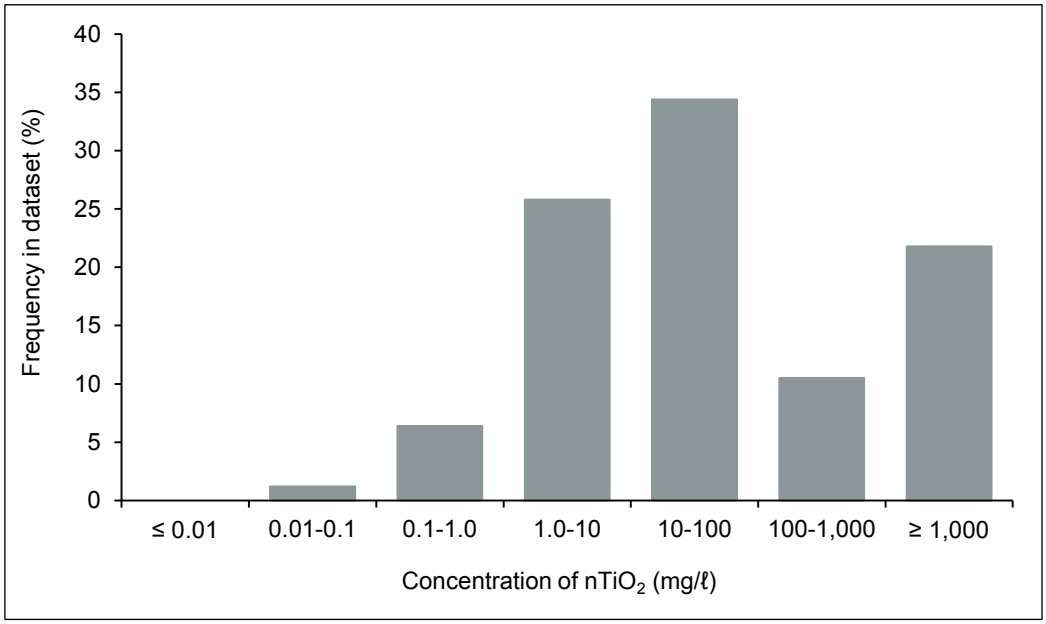


Figure 4.6 Frequency of $nTiO_2$ concentrations in the dataset.

Adsorption studies reported the highest concentration of $nTiO_2$ (Appendix 4.2), inferable to accidental discharges or accumulation in the aquatic environment. Very low concentrations are inferential to attenuations of ENMs in dynamic systems equivalent to reported values in rivers (Neal et al., 2011) (refer to Section 4.2.1). The quality of data based on concentrations of ENMs is 2.88 or 41.1% of the maximum expectation.

Organism culture exhibits 71.3% of the expected quality (Table 4.11), a high score mainly contributed by data generated using axenic cultures of algae species and mixed cultures comprising other organisms. However, all cultures for *D. magna* were axenic and atypical of the aquatic environment. Table 4.12 summarizes descriptor and response variables based on compiled information described in Sections 4.4-4.7.

Table 4.12 Summary of descriptors influencing behaviour kinetics of $nTiO_2$ and toxicity to algae

Details of input variables				Adsorption	Aggregation	Inhibition effects to organisms	
No.	Input variable	Category	Data type	Adsorbed mass (mg/g)	Hydrodynamic size (nm)	Biomass growth (%)	*D. magna* reproduction (%)
1	Composition of anatase crystals (%)	ENM	Continuous	X	X	X	X
2	Charged ions	Abiotic	Categorical	X	X	X	X
3	Organism culture	Biota	Categorical	-	-	X	-
4	Duration (hours)	Abiotic	Continuous	X	X	X	X
5	Ionic strength (mM)	Abiotic	Continuous	X	X	X	X
6	NOM[1]/TC[2] (mg/ℓ)	Abiotic	Continuous	X[1]	X[1]	X[2]	X[2]
7	Organism growth phase	Biota	Categorical	-	-	-	X
8	Organism feeding	Biota	Categorical	-	-	-	X
9	Organism generation	Biota	Categorical	-	-	-	X
10	Organism species	Biota	Categorical	-	-	X	-
11	pH	Abiotic	Continuous	X	X	X	X
12	Surface area concentration (m^2/ℓ)	ENM	Continuous	X	X	X	X
13	Surface coating	ENM	Categorical	-	X	X	X
14	Temperature (^oC)	Abiotic	Continuous	X	X	X	X
15	Types of organic matter	Abiotic	Continuous	X	X	X	-

4.8 Assumptions, Uncertainties, and Limitations

Three paradigms drive modelling assumptions in this study namely, the status of $nTiO_2$ at the time of exposure, and simplicity and applicability of trained models. First, at the time of exposure, intrinsic properties of $nTiO_2$ utilized by industry are fixed. Uncoated pristine $nTiO_2$ used in studies denote ENMs likely released during production and manufacturing stages, as well as those released from nanoproducts during use and disposal stages.

Second, knowledge about physicochemical properties of various types and nanostructures of ENMs, and abiotic factors can allow prediction of potential interaction and toxicity effects in any aquatic system. The pH and temperature are easy to measure and detect in grab samples while ionic strength indicates electrolytes present in water. These variables represent water conditions for the purpose of simple risk assessment exercises, especially where analytical equipment are lacking. Duration of exposure projects retention time used to monitor the persistence of ENMs in the event of aquatic releases.

The data and exposure systems used to generate them, and surface area concentration (SAC) variable may contribute to likely model uncertainties. The modelling approach utilizes statistically averaged secondary data generated using diverse experimental designs and exposure conditions that are not comparable. For example, representing size distribution of particles and aggregates of ENMs in water with a single average value enhances uncertainty likely propagated to model prediction. It is possible to use SAC at the time of exposure in steady systems, but challenging in unsteady aquatic environment owing to speciation and subsequent partitioning of ENMs that may require distribution ratios to determine pathways.

Surface charge density of uncoated ENMs varies with exposure conditions. The substitution of quantitative surface energy values with qualitative surface coating materials as inputs obscure effects of the former, which may lead to prediction uncertainties.

Batch reactors simulate discrete time-controlled processes characterizing lentic aquatic environment and treatment plants. The models developed will be limited to predicting kinetic behavior and toxicity effects of $nTiO_2$ in the semi-static aqueous environment, for example, marshes, pools, ponds, lakes, and water treatment plants.

Biotic factors not related to pollutant effects, for example, resource competition and predation affect organisms, and the presence of other chemical pollutants may enhance or antagonize effects of specific pollutant substances in the aquatic environment. These

scenarios project effects of ENMs that would be different from those observed in pure cultures of the simulated aquatic environment. Hence, toxicity-based models developed in this study will be limited to hypothesis testing for inference with improvement to heterogeneous and dynamic aquatic environment.

4.9 Summary

This Chapter has answered the first research question and objective by establishing inherent physicochemical properties of ENMs, and abiotic and biotic factors influencing behaviour kinetics and biological toxicity during exposure in the simulated aquatic environment. The study critically examined scientific reports on behaviour kinetics and toxicity of $nTiO_2$ as model ENM. Historical information, mathematical models, and statistical analyses aided to reduce data ambiguities that supported collation and construction of a labelled dataset sufficient for data learning.

Five intrinsic physicochemical properties of ENMs, seven abiotic and six biotic factors found applicable to be continuous and categorical input variables for training. Merging three physiochemical properties and two biotic factors to obtain two different inputs reduced the variables from 18 to 15. The database displayed discontinuity between behaviour kinetics and toxicity effects. Therefore, a constructed sample learning dataset constitutes 15 (eight continuous and seven categorical) input variables, and four independent target responses, two from behaviour kinetics of ENMs and two from the toxicity of ENMs to algae and *D. magna* organisms. Based on the criteria developed, the quality of data for use in this study is 1.71 or 44.7% of expected excellence.

Chapter 5 - Prediction of Organic Adsorbates on $nTiO_2$ in Water

5.1 Summary

For the first time, this Chapter reports the application of multilayer perceptron neural networks (MLP-NN) model to predict adsorbed mass of organic matter on nano-titanium dioxide ($nTiO_2$), a model engineered nanomaterial (ENM). The model was implemented using two properties of the ENM (anatase composition in $nTiO_2$ crystals and surface area concentration (SAC)); and seven abiotic factors (pH, ionic strength, temperature, duration of exposure, types and concentration of organic matter, and charged ions).

An implemented ensemble of 25-standalone models exhibited average training and cross-validation set errors of 0.058 ± 0.002 and 0.082 ± 0.012, respectively. The predictive capability of the implemented ensemble model displayed a mean square error of 0.101 and $R^2 = 0.931$, demonstrating good generalization capability. Input perturbation and ranking methods established the temperature > duration of exposure > types of organic matter > composition of anatase $nTiO_2$ crystals > SAC > concentration of organic matter > pH > charged ions > ionic strength as the sequence of influential input variables. By replacing the mean with median as perturbed input data, reordered and consistent ranking of the first six and last three inputs, respectively, indicate marginal influence of pH, charged ions, and ionic strength on adsorption.

Model findings reveal potential under- or overestimation of natural organic matter adsorbates on $nTiO_2$ in the aquatic environment by use of surrogate organic matter and electrolytes in generic water systems. Contrary to published literature, the influence of K^+ on adsorption appears greater than Ca^{2+} and Mg^{2+} in acidic conditions and Ca^{2+} in basic conditions. The results provide a basis for using the MLP-NN as a tool for future development of dynamic models in predicting adsorption kinetics involving ENMs in the aquatic environment.

5.2 Introduction

After their release into the aquatic environment, ENMs would interact with organic and inorganic aquatic ligands. Chapter 4 discusses in-depth factors influencing adsorption not repeated in this Section. Modelling studies have not investigated the adsorption of organic matter on ENMs using the multilayer perceptron neural networks (MLP-NN) model. Thus,

this study relates model findings to the existing body of knowledge on adsorption in published literature. The terms ‘mass coverage of organic matter or ligands or adsorbates’ and ‘adsorbed mass’ interchangeably refer to a number of molecules adsorbed on ENMs’ surface. The ‘adsorption capacity’ refers to the maximum coverage of adsorbates on ENMs.

Several experimental studies have investigated the adsorption of various types of organic matter on $nTiO_2$, tested selected causal factors, and in some instances isotherm models fitted to experimental data to explain equilibrium buildup of adsorbates. Section 4.6.1 in Chapter 4 describes thermodynamic models applied. This study applies data collated from some of these independent studies’ reports presented in Section 4.6.1, Table 4.3, and Appendix 4.2, from which nine causal factors are model inputs.

Possibilities of applying diverse thermodynamic models to explain adsorption constants and adsorbed mass demonstrate size-related heterogeneous surface characteristics of ENMs that avails multiple binding sites and different bonding mechanisms of organic matter. Barreto et al. (2012) proposed that BET was the most suitable isotherm in explaining adsorption behaviour involving $nTiO_2$. Using carbon nanotube (CNT) as a model ENM, the BET model outperformed the Freundlich and single mode Langmuir isotherms in explaining binding affinities of two polycyclic aromatic hydrocarbons (Kah et al. 2011). However, the performance of the BET was outpaced by dual mode and dual Langmuir isotherms, as well as the Dubinin-Ashtakhov Polanyi theory-based model (Kah et al. 2011), a finding reported elsewhere for different organic compounds (Yang and Xing 2010). Therefore, applicable isotherms depend on properties of adsorptive, adsorbent ENMs and exposure conditions, from which model input variables applied in this study were drawn.

Section 4.4 in Chapter 4 describes in details the SAC and composition of anatase crystals representing the adsorbent in this study. Moreover, the isotherm models estimate adsorbed mass using concentration of adsorptive and adsorbent as inputs. Hence, the concentration of organic matter and SAC property of ENMs are relevant inputs in modelling adsorption kinetics. Training data comprise bare (uncoated) ENMs as the only nominal level of surface coating; therefore, this variable is left out for lack of statistical variation.

Notably, equilibrium states of adsorptive and adsorbates were measured at a constant narrow range of values for the temperature (20 to 25 ^{o}C) and duration of exposure (0.5 to 48-hours). The temperature affects solubility and adsorption of organic moieties while the duration of exposure regulates the transfer rates of adsorptive, interaction and partitioning on solid interfaces. The adsorption-desorption enthalpies and entropy of polar and non-polar

adsorptive, respectively, depend on temperature (Hamaker and Thompson 1972). The effects of interaction time on adsorption capacity reduces with increasing concentration of adsorptive (Elimelech et al., 1995). Thus, investigating their influence on adsorption involving ENMs will broaden an understanding that is inferable to the aquatic environment.

The adsorptive applied in this study comprise ten nominal levels of natural organic matter (NOM) and surrogate organic matter (SOM) listed in Table 4.7 in Chapter 4. The origin of natural organic matter (NOM) and Suwanee River NOM (SRNOM) distinguish their specific properties. The SRNOM, Suwanee River fulvic acid (SRFA), Suwanee River humic acid (SRHA), oxalic, and citric acids are five single adsorptive molecules. The 2-(N-morpholino) ethanesulfonic (MES) and 4-2(-hydroxyethyl)-1-piperazineethanesulfonic (HEPES) buffers used alongside adsorption of adipic, citric and oxalic acids (Pettibone et al. 2008; Mudunkotuwa and Grassian 2010) necessitated combining four dual-organic acid adsorptive molecules (refer to Table 4.7). The competition of buffers with humic acid on $nTiO_2$ (Erhayem and Sohn 2014a) and temperature effects on MES and HEPES as well as their adsorption behaviour (Kandegedara and Rorabacher 1999; Mash et al. 2003; Sokołowska and Bal 2005) motivates their inclusion as dual-organic acid adsorptive. This study applies the MLP-NN model to predict mass coverage of these adsorptive on ENMs under solid-liquid interfacial conditions.

Majority research studies have investigated the effects of mainly single monatomic mono- and polyvalent ions on adsorption. However, speciation of mixed mono-, poly- and multivalent ions may affect adsorption and partitioning of ENMs in the aquatic environment. In this study, the categorical variable ‘charged ions’ represents identified mono- and divalent ions, and mixtures of identified and unidentified ions found in purified, standard synthetic and natural water. Thus, nine different nominal levels of charged ions investigated include mono ions (K^+, Na^+, Ca_{2+}, and Mg^{2+}) in purified water, and mixed ions in river water from three different locations ($RW_{1\text{-}3}$) (Chekli et al. 2015), seawater (SW) (Chekli et al. 2015), and synthetic moderate hard water (MHW) (USEPA 2002). Unknown ions designate those not determined by the design of experiments that generated the data used.

Results discussed in this Chapter address the first part of the second objective to answer part one of the second research question, which sought to establish relationships between properties of ENMs and environmental abiotic and biotic factors, and adsorption kinetics involving ENMs. Section 3.2 of Chapter 3 describes methods applied to implement the ensemble paradigm of neural networks to approximate the response data. Using the fitted

neural network model to evaluate inputs and response sensitivities establishes the relative importance of variables and hypothetical input-output relationship patterns generated using interpolated data. Application of study findings relates to mobility, bioavailability, toxicity, and management of ENMs in the aquatic environment.

5.3 Results and Discussion

5.3.1 Network Design, Training, and Aggregation of Models

Implementing procedures outlined in Section 3.2.1 assigned fractional numerical codes to categorical data; normalized inputs, except coded values, to lie between -0.9 and 0.9; and standardized outputs to have a mean of zero and variance of one. Application of Equation 3.5 in Chapter 3 determined the number of hidden neurons between 3 and 8, inclusive, totaling 6 different network configurations. The tanh and identity were set as transfer functions in the hidden and output layers, respectively. The number of iterations was set at 1,000 to guard against potential model over-fitting.

The data were randomly split into training (264), cross-validation (66) and out-of-sample testing (36) sets based on the method described in Section 3.2.3 in Chapter 3. Initial seeds for random subsampling during parallel training and cross-validation were set at 1, 123, 1,234, 5,678, and 9,999 and 5 models retained after every training experiment. While not a foolproof procedure, subsampling assigned different data values into the two subgroups to improve learning from a heterogeneous dataset. However, similar data values fell into the same allotment and somewhat reduced the intended diversity.

Training preliminary models established 5.5×10^{-5} and 5.5×10^{-6} as suitable weight decay, (α) values (refer to Eqn. 3.13 in Section 3.1.4) for use in regulating the magnitude of connection weights and reducing performance errors in the hidden and output layers, respectively. Thirty five models were trained using the training and validation sets at fixed random seed of 1,234 and 6 hidden neurons, and 7 α values varied in 10-fold increments between 10^{-7} and 10^{-1} for hidden layer, and 10^{-8} and 10^{-2} for output layer (Table 5.1).

The models exhibiting the smallest average validation set error determined the selected α values used to train implemented ensemble model. The use of large and small α values have demonstrated under- or over fitted models, respectively (Murphy 2012), suggesting careful selection of α to achieve a smooth generalizing function.

Table 5.1 Weight regularization values for organic adsorbate-based models

Weight regularization (α)		MSE of standalone models		Ensemble	
Hidden layer	Output layer	Validation	Test	Test MSE	R^2
10^{-7}	10^{-8}	0.110 ± 0.073	12.83 ± 32.05	0.961	0.520
10^{-6}	10^{-7}	0.088 ± 0.067	1.003 ± 1.921	0.197	0.776
10^{-5}	10^{-6}	0.081 ± 0.047	0.927 ± 2.124	0.363	0.866
10^{-4}	10^{-5}	0.082 ± 0.043	0.180 ± 0.170	0.133	0.910
10^{-3}	10^{-4}	0.107 ± 0.037	0.290 ± 0.068	0.266	0.821
10^{-2}	10^{-3}	0.300 ± 0.120	0.854 ± 0.327	0.731	0.571
10^{-1}	10^{-2}	0.540 ± 0.043	1.473 ± 0.031	1.472	0.070

Table 5.2 summarizes the performance of one hundred and fifty standalone models retained, 25 models for each network topology evaluated that form the basis of subsequent discussions. The topology column describes multilayer perceptron (MLP) algorithm used, and the number of model inputs (9), hidden neurons (3 to 8), and output neuron (1).

Table 5.2 Summary of standalone and ensemble model performance

	Standalone model MSE (0.95 confidence interval)			Ensemble model MSE		Ensemble
Topology	Training	Validation	Testing	Testing	Reduction (%)	R^2
MLP 9-3-1	0.066 ± 0.005	0.090 ± 0.012	0.205 ± 0.098	0.130	36.4	0.911
MLP 9-4-1	0.060 ± 0.003	0.085 ± 0.011	0.132 ± 0.023	0.092	30.44	0.938
MLP 9-5-1	0.059 ± 0.003	0.084 ± 0.012	0.149 ± 0.044	0.100	33.7	0.933
MLP 9-6-1	0.058 ± 0.002	0.082 ± 0.012	0.157 ± 0.075	0.101	35.8	0.931
MLP 9-7-1	0.058 ± 0.003	0.083 ± 0.012	0.134 ± 0.044	0.086	35.9	0.942
MLP 9-8-1	0.057 ± 0.002	0.082 ± 0.013	0.133 ± 0.030	0.089	32.9	0.935

The results demonstrate that varying the number of hidden neurons did not significantly influence the average cross-validation set errors (Table 5.2). An increase in the number of hidden neurons increases the degrees of freedom because of increasing connection weights, thus, model complexity. However, the results reflect the likely stability of the BFGS training algorithm used (Necedal and Wright 2006). Despite comparable validation set errors, generalization capability of the models varied and displayed higher testing set

prediction errors (Table 5.2). The finding suggests that models selected based on the smallest validation set errors may not certainly generalize well on new data. Figure 5.1 illustrates the graphical error variances of the validation set (*Top*) and out-of-sample testing set (*Bottom*). Moreover, the improvement in average prediction errors and R^2 that show good approximation of experimental data agrees with findings of Naftaly et al. (1997) that showed reduced errors of aggregated models trained by varying initial seeds.

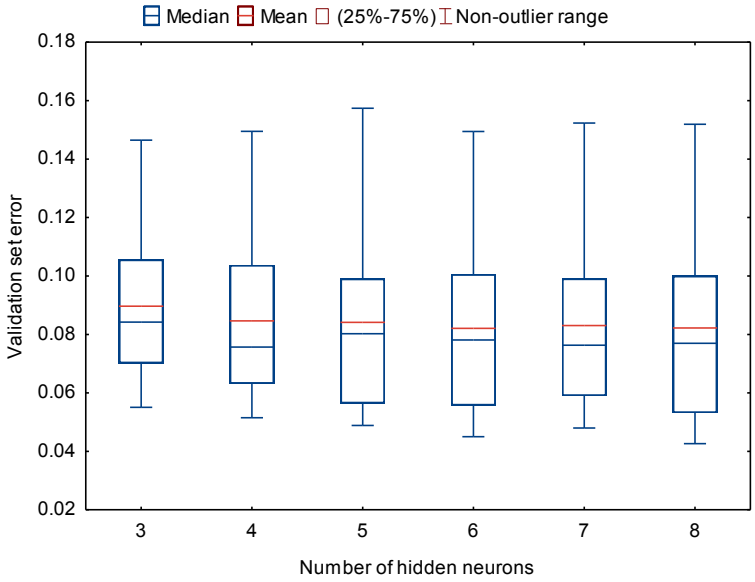
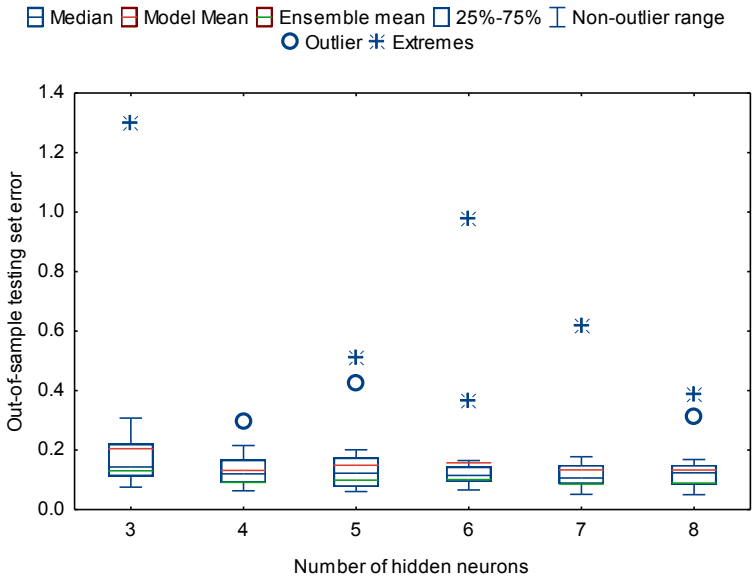


Figure 5.1 Average validation (*Top*), and testing (*Down*) set errors of networks having a different number of hidden neurons.



All models presented in Table 5.2 adequately approximated adsorbed mass by using different populations of training examples and generated diverse error variances based on network topologies. The testing set error is not a measure of selecting models to implement, and this study considers aggregating models having a similar number of hidden neurons to reduce subjective optimism. Thus, the minimum validation set error was used as a criterion to select and implement (Murphy 2012) the MLP 9-6-1 ensemble model (Table 5.2) (in grey).

The performance error of the selected ensemble model based on 90% training and validation sets was 0.116. Applying the out-of-sample test set data, the implemented model adequately predicted mass coverage of organic matter on $nTiO_2$ with an $R^2 = 0.931$ (0.95 prediction interval) (Table 5.2 and Fig. 5.2).

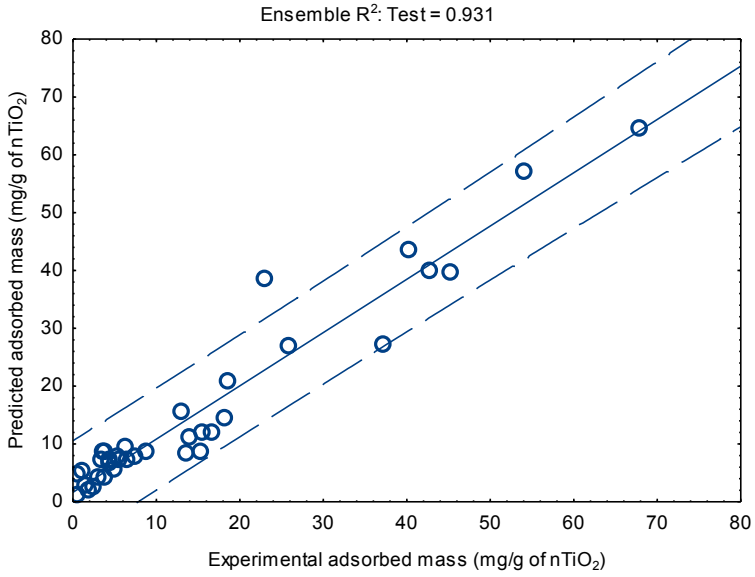


Figure 5.2 Ensemble prediction of mass coverage of organic matter on $nTiO_2$.

A scatter comparison of predicted versus residual adsorbed mass reveals model independence with $R^2 = 0.001$ and slope = 0.002 (Fig. 5.3) (*Left*). The normality of residuals is high $(W = 0.914)$ but a positive (right-skewed) distribution describes the deviation of the fitted and actual adsorbed mass values (Fig. 5.3) (*Right)*. The residuals' mean, median, minimum, maximum and standard deviation values are -0.452, -0.864, -15.64, 10.16 and 4.42, respectively. Eighty-seven percent residuals are spread between -5 and 5 with two outliers outside the 95.0% prediction interval. However, 67.0% negative residuals demonstrate mostly over-predicted adsorbed mass with majority small deviations (Fig. 5.3) (*Right*)

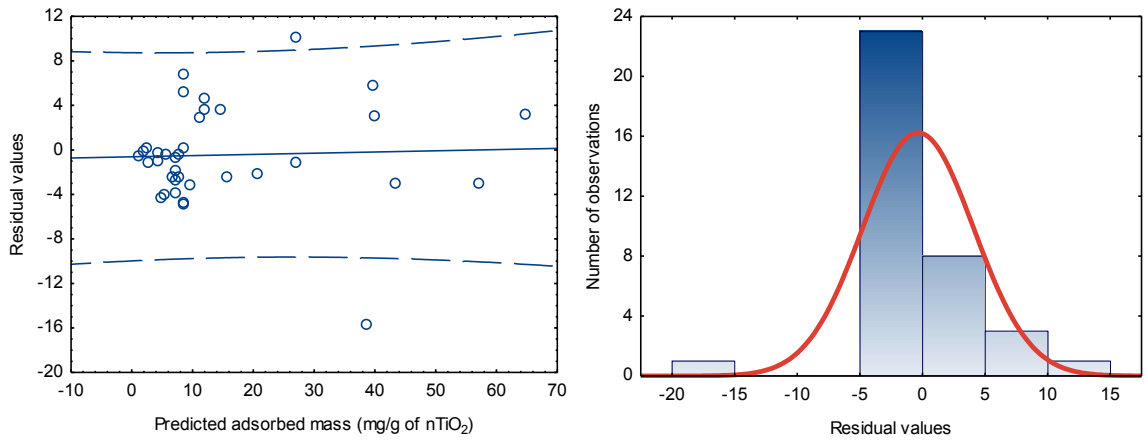


Figure 5.3 Predicted versus residual adsorbed mass (*Left*) and normality of residual values (*Right*).

5.3.2 Relative Importance of Input Variables

The relative importance of variables affecting the quantity of organic adsorbates on $nTiO_2$ was evaluated by way of input perturbation and ranking method described by Francis (2001) and discussed in Section 3.2.4. Table 5.3 presents errors computed by perturbing input variables using their mean values in the ensemble model and corresponding sensitivities when compared to the error of the full model. The assigned ordinal numbers in descending order of rank shows the direction of importance. The most important variable displays high MSE / sensitivity when perturbed and vice versa.

Table 5.3 Relative importance of input variables in predicting adsorbed mass on $nTiO_2$

Input variable	MSE	Sensitivity (%)	Rank
Temperature (°C)	1.332	91.3	1
Duration of exposure (hours)	0.963	88.0	2
Types of organic matter	0.937	87.7	3
Composition of anatase crystals (%)	0.855	86.5	4
Surface area concentration (SAC) (m^2/ℓ)	0.732	84.2	5
Initial concentration of organic matter (mg/ℓ)	0.353	67.2	6
pH	0.248	53.5	7
Charged ions	0.164	29.9	8
Ionic strength (mM)	0.118	1.9	9

The sensitivity analysis reveals temperature to be the strongest predictor variable though its effects on adsorption have yet to be evaluated while ionic strength is the least influential predictor. The temperature determines the thermal stability of organic molecules, hence, their hydrophobic and hydrophilic partitioning with the former easily removed from the aqueous solution (Site 2001). In relation to the duration of exposure, the temperature determines adsorption rates (Hulscher and Cornelissen 1996) of inner and outer sphere attachment.

The crystal structure of $nTiO_2$ appears to be a stronger predictor than the SAC. Whereas SAC may avail reactive surface sites for adsorption, the crystal structure and stability under diverse conditions (Diebold 2003) determine unsaturated or stable coordination of $nTiO_2$ atoms at the available surface sites to influence reactivity. Moreover, the types of organic matter is more influential than the initial concentration of organic matter, which implies the possibility of molecular structure and strength of interfacial bonds determining the quantity of adsorbed or desorbed organic matter at adsorption sites.

Notably, the reported sensitivities, thus, relative importance, are specific for the selected ensemble model because different associations of input variables approximate adsorbed mass of organic matter on $nTiO_2$. Therefore, an evaluation of input variables' combinations based on other ensemble models presented in Table 5.2, or their covariance, would yield different relative importance ranks. Thus, to assess the consistency of the ranks presented in Table 5.3, the median values replaced the mean values of continuous variables in the fitted ensemble model. A comparison between the new prediction errors and the error of the full model revealed a revised order of variable significance, viz. duration of exposure > temperature > types of organic matter > SAC > composition of anatase crystals > concentration of organic matter > pH > charged ions > ionic strength.

Coincidentally, the sensitivity of pH and ionic strength did not change whilst that of the concentration of organic matter increased by 38.5%. The last four variables listed in Table 5.3 retained their ranks, duration of exposure and temperature interchanged positions, and SAC replaced composition of anatase crystals, which dropped one place downwards. Moreover, ranks for types of organic matter and charged ions did not change. Concerning the application of the ensemble model, the consistency tests imply low influence of pH and ionic strength on adsorption and suitability of neural networks in forecasting adsorption outputs based on different combinations of important input variables.

To discern the positive or negative effects of the various types of organic matter and charged ions on adsorption based on individual substitution in the ensemble model may be difficult. However, the individual sensitivities of each nominal level of organic matter and charged ions listed in Table 5.4 connote two key issues; first, adsorption bonding mechanisms differ with the properties of organic molecules. Second, mixed ions found in natural water influence adsorption more than mono- and divalent ions spiked in purified and standard synthetic water. The controlled salt concentrations in MHW water limits speciation of ions, which may explain low sensitivity compared to natural water despite comprising mixed ions.

Table 5.4 Relative importance of diverse organic matter and charged ions in predicting adsorbed mass on $nTiO_2$

Types of organic matter	MSE	Sensitivity	Rank	Charged ions	MSE	Sensitivity	Rank
NOM	2.466	95.31	1	$RW_{3(Mixed)}$	0.189	38.96	1
SRHA	2.031	94.31	2	$RW_{2(Mixed)}$	0.187	38.14	2
SRNOM	1.518	92.40	3	$RW_{1(Mixed)}$	0.182	36.53	3
SRFA	0.883	86.91	4	Mg^{2+}	0.172	33.00	4
HEPES Citric acid	0.534	78.34	5	K^{+}	0.159	27.44	5
Oxalic acid	0.488	76.30	6	Ca^{2+}	0.150	23.16	6
MES Citric acid	0.483	76.06	7	$MHW_{(Mixed)}$	0.146	20.62	7
Citric acid	0.408	71.70	8	Na^{+}	0.122	5.00	8
HEPES Oxalic acid	0.321	64.01	9	-	-		-
MES Adipic acid	0.240	51.87	10	-	-		-

The SOM are subcomponents of complex NOM (Buffle et al. 1998), whose preferential adsorption mechanisms (Gu et al. 1995) may explain the high sensitivity. In agreement with findings by Erhayem and Sohn (2014b), assorted molecular properties and concentration of polar and non-polar fractions affected adsorption in the sequence SRHA > SRNOM > SRFA (Table 5.4). The sensitivities imply SRHA and SRNOM surrogates are closest to NOM and applicable in predicting adsorption of organic matter on ENMs in the aquatic environment. Application of SRFA, carboxylic and dual buffer-carboxylic acids in modelling would underestimate NOM adsorption with varying margins of error.

Pettibone et al. (2008) remarked about unlikely effects of buffers to adsorption. However, mixtures of HEPES and MES with organic acid molecules reveal the potential low or high

adsorption influence of the buffers, with lower effects of MES possibly caused by weak binding abilities (Mash et al. 2003). The HEPES oxalic acid exhibits low importance compared to oxalic acid, implying possible competition of the buffer with the acid on $nTiO_2$. The high ranking of HEPES and MES citric acids compared to citric acid connotes higher affinity of dual-organic molecules related to the presence of buffers. The MES adipic acid portrays the least influence on adsorption compared to the other single or dual organic acid, probably because of its monodentate binding mechanism.

In natural water, the sequence $RW_3 > RW_2 > RW_1$ (Chekli et al. 2015) shows likely influence of concentration and speciation of free ions on adsorption. In published literature, decreasing valence of exchangeable cations reduces adsorption of organic molecules on metal oxides owing to proton competition and co-adsorption with monovalent and divalent ions, respectively. The divalent ions, especially Ca^{2+}, are associated with bridging interaction with functional groups and ENMs (Chen and Elimelech 2008; Sun and Lee 2012; Chowdhury et al. 2012; Erhayem and Sohn 2014a; Chekli et al. 2015). Assuming adsorption effects, (Von Der Kammer et al. 2010) remarked about the sole influence of Ca^{2+} on ENMs' behaviour in the aquatic environment. However, the significance of $Mg^{2+} > K^{+} > Ca^{2+} > Na^{+}$ (Table 5.4) contrasts findings reported in the published literature.

The low influence of Ca^{2+} compared with K^{+} may be attributable to their cation exchange capability with different adsorptive under similar exposure conditions. Thus, it is possible that effects of organic and inorganic adsorption will interfere with Ca^{2+} speciation. Moreover, proton complexation in acidic conditions (Murphy et al. 1994), absent Ca^{2+}-organic matter complexes in basic conditions (Davis 1982), and influence of ionic size and cation-specific binding mechanisms on inter-particle attraction and bridging of monovalent ions (Guerrero-Garca et al. 2011; Wang et al. 2011) may reduce effects of divalent cation bridging.

It is unexpected that studies attribute ionic strength to adsorption (Erhayem and Sohn 2014b; Chekli, Zhao, et al. 2015; Chekli, Roy, et al. 2015) but the variable is the least influential. Davis (1982) recorded none and minimal effects of ionic strength on adsorption of NOM on metal oxides in monovalent (Na^{+}) < 10 mM and > 100 mM regimes, respectively. Thus, the marginal significance ascertained by the model likely relates to the data used - 87.0% generated in monovalent regimes with 99.0% of ionic strength ≤ 100 mM.

5.3.3 Simulated Explanatory Input-Adsorbed Mass Relationships

The method proposed by Lek et al., (1996), described in detail in Section 3.2.4 of Chapter 3, was used to evaluate the explanatory contribution of input variables in predicting organic adsorbates on $nTiO_2$. The statistically sorted data simulates five hypothetical exposure conditions represented by the minimum, lower quartile (LQ), median, upper quartile (UQ) and maximum data values of input variables in the dataset (Table 5.5). The adsorptive evaluated are SRHA, SRFA, MES adipic acid, HEPES oxalic acid, and NOM (Table 5.5).

Table 5.5 Hypothetical input data used to predict adsorbed mass of organic matter on $nTiO_2$

Input variable	Minimum	LQ	Median	UQ	Maximum
SAC (m^2/ℓ)	0.25	0.75	164.00	876.00	876.00
Temperature (°C)	20	22	25.00	25.00	25.00
Concentration of organic matter (mg/ℓ)	1.01	10.01	23.74	90.98	917.08
Duration of exposure (hours)	0.50	2.00	2.00	48.00	48.00
Types of organic matter	SRHA	SRFA	MES adipic acid	HEPES oxalic acid	NOM
Charged ions	K^+	Na^+	Na^+	Na^+	RW_3
Composition of anatase crystals (%)	80.0	82.0	100.0	100.0	100.0
pH	2.00	4.80	5.50	7.50	8.04
Ionic strength (mM)	4.00	20.00	20.00	30.00	344.00

Restricted to using the same range of values for continuous variables, it is conceivable to estimate 90 response patterns while changing nominal levels listed in Table 5.5. However, likely complications in interpreting findings limited exploration of this aspect. Instead, interval values used for each hypothetical exposure condition (Table 5.5) yielded 60 different predicted values of adsorbed mass. Applying the numerical codes for 10 and 8 nominal levels of types of organic matter and charged ions estimated 50 and 40 response outputs, respectively.

Lek et al. (1996) suggested using the median of predicted response outputs to simulate trends and compute relative importance of input variables. However, this study discusses independent trends generated by the minimum, LQ, median, UQ and maximum data values to evaluate explanatory contribution of input variables to predicting adsorbed mass without ranking the variables. Thus, five independent profiles were plotted for continuous variables,

and histograms for nominal levels of categorical variables versus projected adsorbed mass. The large increments in values of SAC, ionic strength and initial concentration of organic matter brought about small changes in simulated adsorbed mass outputs. Also, a similar effect was caused by small increments in composition of anatase crystals, pH, temperature and duration of exposure.

Both positive and negative values of adsorbed mass were estimated, which suggest an increase and decrease in adsorption with respect to the input variables, respectively. Since multiple factors contribute to the estimated response outputs, it may be difficult to precisely explain the individual trends generated. Hereafter, deliberations underscore the implications of variable importance on adsorption, regarding Tables 5.3, 5.4 and 5.5 and published literature on adsorption.

a) Temperature

The adsorption is high under low temperatures and increases when the temperature increases in dissimilar trends for SRHA, SRFA, MES adipic acid, HEPES oxalic acid, and NOM (Fig. 5.4). Adsorption of HEPES oxalic acid decreases after the temperature ≥ 23.0°C.

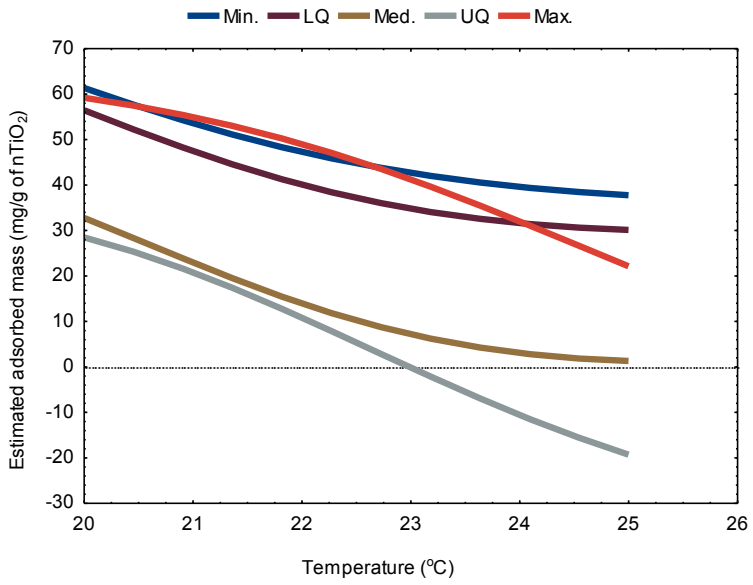


Figure 5.4 Profiles of adsorbed mass estimated by varying the temperature.

An increase in temperature is known to influence concurrent fast adsorption and desorption under equilibrium enthalpies (Hulscher and Cornelissen 1996), solubility of polar components of adsorptive (Chiou et al. 1979) or exothermic adsorption kinetics (Site 2001; Wang et al. 2012), which decrease adsorbed quantities. It is likely that MES adipic acid and

HEPES oxalic acid form weaker bonds with $nTiO_2$ as temperature increases. Further investigations are needed to corroborate the projected trends.

b) Duration of Exposure

The adsorption is high for SRHA, SRFA, and NOM, low for HEPES oxalic acid and very low for MES adipic acid at short exposure periods (Fig. 5.5). Whereas the adsorption of SRHA, SRFA, and NOM drops, HEPES oxalic and MES adipic acids' adsorbates decrease with increasing duration of exposure. The equilibrium adsorption of organic matter on $nTiO_2$ decreased with exposure periods (Erhayem and Sohn 2014b) likely because of rate-limiting effects of lengthy exposures (Site 2001). There is likely high surface reaction and ligand exchange at adsorption sites at short exposure periods. With lengthy exposures, the maximum surface coverage of adsorptive and transformation of adsorbents surface chemistry likely reduces or hinders adsorption, respectively, thus decreasing adsorbates.

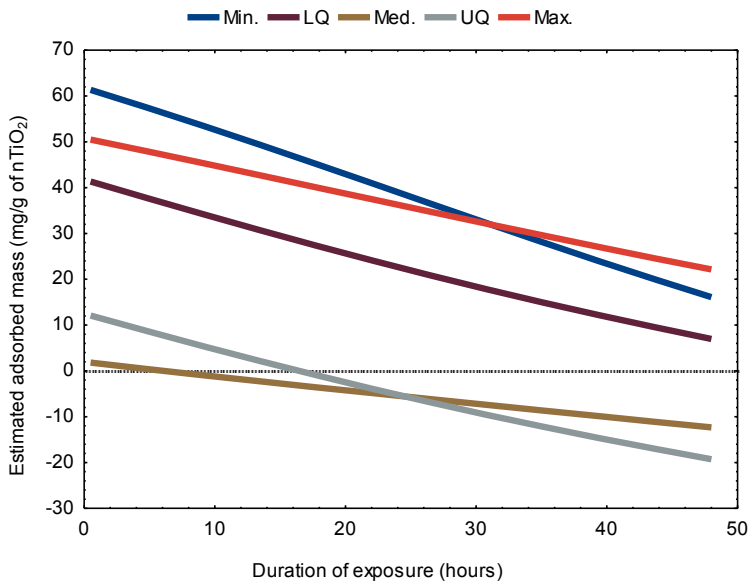


Figure 5.5 Profiles of adsorbed mass estimated by varying the duration of exposure.

c) Types of Organic Matter

The extracted trends suggest higher effects of minimum-LQ data values on adsorption than median-to-maximum data values. Under all simulated exposure conditions, the NOM adsorbates were positive and higher that the SOM adsorbates (Fig. 5.6), which confirms the leading sensitivity of the former (Table 5.4) but varied influential order for the latter types of organic matter. The trends confirm the influence of adsorptive properties (Erhayem and Sohn 2014a) and binding mechanisms (Philippe and Schaumann 2014) on adsorption.

The minimum data estimated the highest adsorbed mass for all adsorptive in the sequence NOM > oxalic acid > HEPES oxalic acid > MES adipic acid > citric acid > MES citric acid > HEPES citric acid > SRFA > SRNOM > SRHA, followed by LQ data in the same sequence. These conditions are likely to promote strong binding of negatively charged functional groups and positively charged $nTiO_2$ (Yang et al. 2009; Mwaanga et al. 2014).

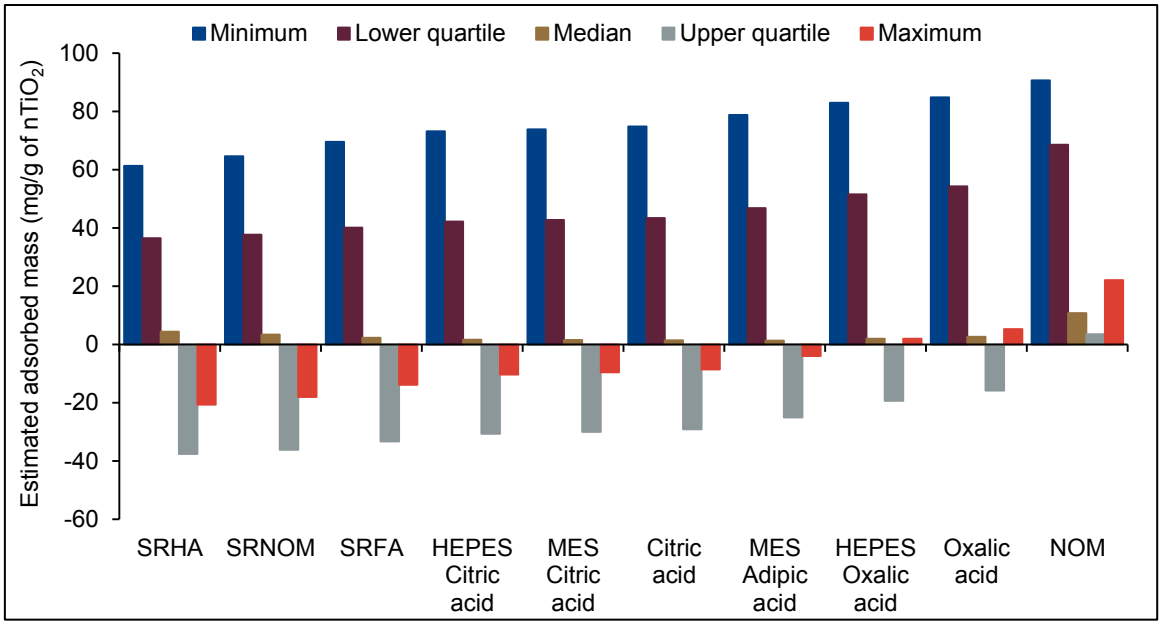


Figure 5.6 Profiles of adsorbed mass estimated by varying the types of organic matter.

The median data values estimated low positive adsorbates in the sequence NOM > SRHA > SRNOM > Oxalic acid > SRFA > HEPES oxalic acid > HEPES citric acid > MES citric acid > citric acid > HEPES adipic acid. The data simulates a transitioning acid-to-basic condition that favours $nTiO_2$ clustering having low net surface charge, hence, weak electrostatic interaction of the adsorptive with the absorbent.

Multiple interacting factors likely influence adsorption under conditions simulated by UQ and maximum data. The UQ data estimated decreasing SRHA < SRFA < SRNOM < HEPES citric acid < MES citric acid < citric acid < MES adipic acid < HEPES oxalic acid < oxalic acid except for low NOM adsorbates. Similarly, the maximum data projected decreasing adsorbates except for low quantities of adsorbed HEPES oxalic acid < oxalic acid < NOM.

d) Composition of Anatase Crystal Phase

The high adsorption of SRHA, SRFA, NOM, and MES adipic acid drops while that of HEPES oxalic acid decreases with increasing composition of anatase $nTiO_2$ crystals (Fig. 5.7). The

anatase coexists with rutile, and occasionally brookite, phases (refer to Section 4.4.1). The trends suggest higher combined adsorbates on mixed than on anatase $nTiO_2$ crystals, likely owing to different crystallographic configurations and surface adsorption mechanisms.

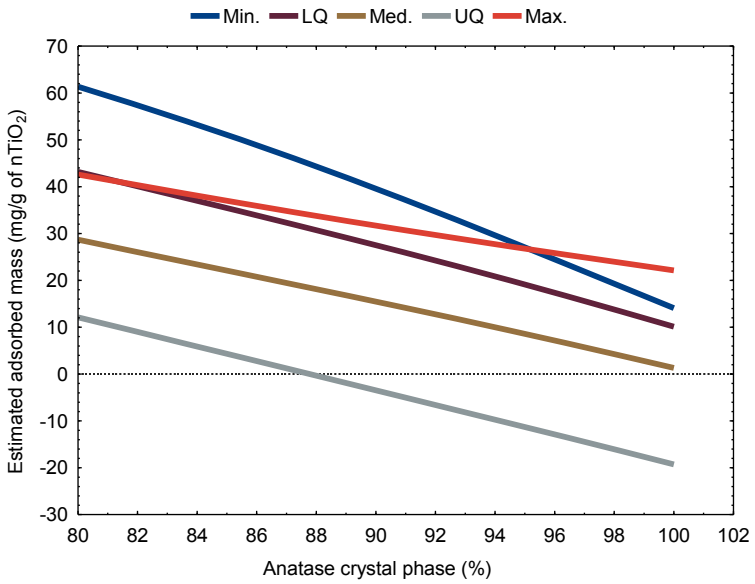


Figure 5.7 Profiles of adsorbed mass estimated by varying the composition of anatase $nTiO_2$ crystals.

The atomic arrangement of anatase and rutile crystals, and unique speciation of surface moieties determine stable and unstable bonding mechanisms with oxalate - with unstable tridentate observed on anatase (Mendive et al. 2009). Conversely, the UQ pattern contrasts theoretical computations by Fahmi et al., (1995) as cited by (Diebold 2003) and experimental findings by Pettibone et al. (2008) that revealed strong binding of oxalate on *ca.* 100.0% anatase than on *ca.* 100.0% rutile owing to a high surface reactivity of the former. Thus, the theoretical notions from these trends require empirical determination for validation.

e) Surface Area Concentration

The increase in adsorption with increasing SAC is more evident for SRHA and SRFA than NOM, the latter decreasing at low SAC then increasing afterward (Fig. 5.8). Moreover, the adsorption of MES adipic acid appears complex, showing a low increase, decrease and increase at low, intermediate, and high SAC values, respectively. In a similar pattern to MES adipic acid, adsorption of HEPES oxalic acid decreases with increasing SAC.

Independently, the primary size and concentration of $nTiO_2$ have been linked to increasing or decreasing adsorption of organic matter on the nanoparticle depending on exposure conditions (Zhang et al. 1999; Von Der Kammer et al. 2010; Roncaroli and Blesa 2011; Sun

and Lee 2012; Chowdhury et al. 2013). The inference from experimental findings is not applicable to SAC. However, the MLP-NN model provides insight on the potential application of SAC as a promising predictor of adsorption of organic matter on ENMs.

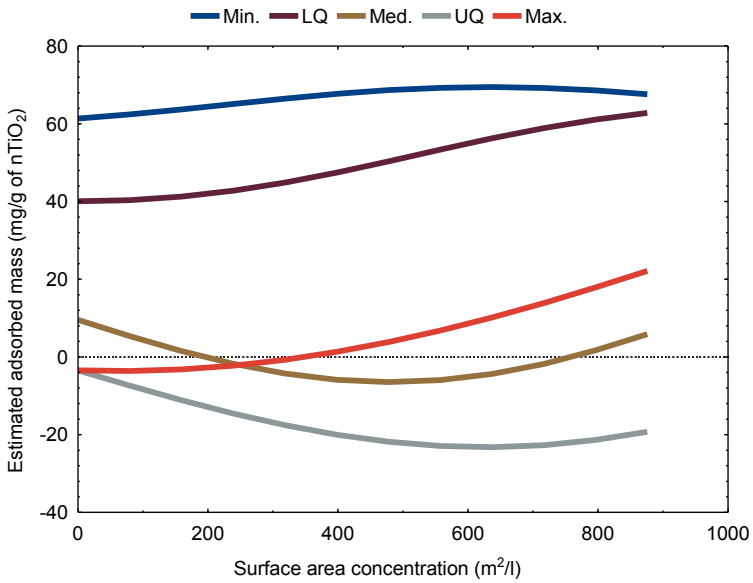


Figure 5.8 Profiles of adsorbed mass estimated by varying the surface area concentration.

f) Initial Concentration of Organic Matter

There is maximal concentration of SRHA, SRFA, and MES adipic acid where adsorption is high before declining (Fig. 5.9).

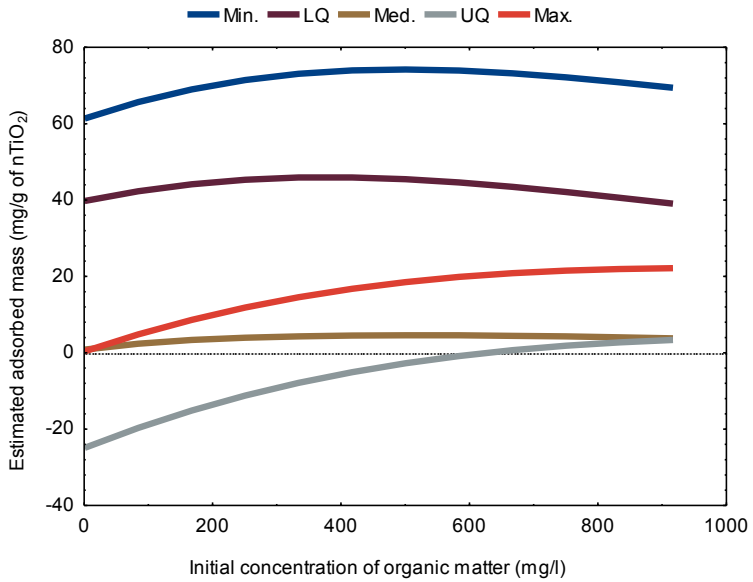


Figure 5.9 Profiles of adsorbed mass estimated by varying the concentration of organic matter.

Conversely, the adsorption of NOM and HEPES oxalic acid increases with increasing concentration, the latter increasing after *ca.* 630-hours. The simulated patterns are in agreement with study findings that attributed an increase in adsorption of SRHA (Sun and Lee 2012), SRFA (Domingos et al. 2009), adipic and oxalic acids (Zhang et al. 1999; Pettibone et al. 2008) and NOM (Chekli et al. 2015; Chekli et al. 2015) to increasing initial concentration of the adsorptive.

g) *pH*

The adsorption of SRHA, SRFA, and NOM drops, marginally increases for HEPES oxalic acid and decreases for MES adipic acid with increasing pH (Fig. 5.10). At low pH, strong electrostatic forces and affinity for conjugate bases by positively charged $nTiO_2$ enhance adsorption (Sun and Lee 2012; Chowdhury et al. 2012). At high pH, dissociation of functional groups that remain in solution, electrostatic repulsion of negatively charged adsorptive and adsorbent, and reduced adsorption sites reduce adsorption(Tipping 1981; Yang et al. 2009). Thus, the trends for SRHA and SRFA agree with results (Zhou et al. 2013; Erhayem and Sohn 2014b) and for MES adipic acid marginally parallel findings (Pettibone et al. 2008).

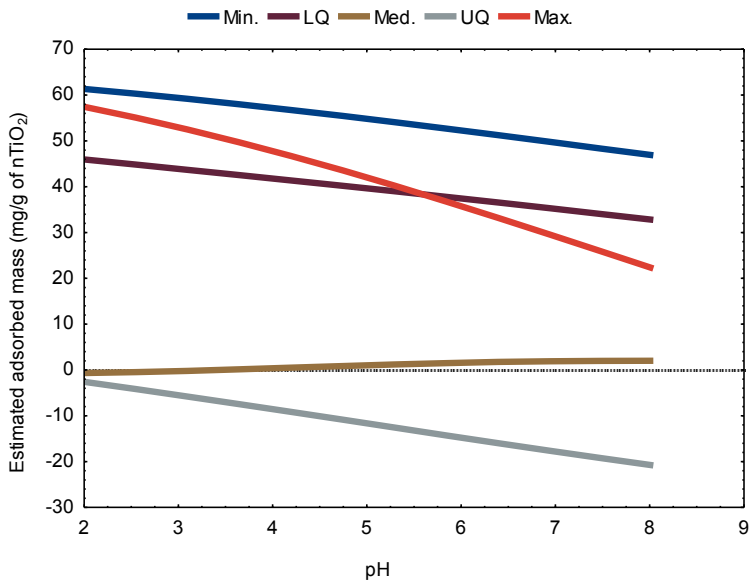


Figure 5.10 Profiles of adsorbed mass estimated by varying the pH.

However, the NOM trend contrasts results (Chekli et al. 2015; Chekli et al. 2015) that did not relate adsorbed mass to pH in natural water. The non-specific electrophoretic mobility of $nTiO_2$ in natural water (Keller et al. 2010) suggests multifarious interactions and preferential adsorption of NOM (Gu et al. 1995) leading to a decrease in adsorbates with increasing pH.

Additionally, the pH range evaluated goes beyond the optimum (pH 4) for equilibrium adsorption of oxalic acid (Mazellier and Sulzberger 2001; Zhang et al.,1985 as cited by Lan et al. 2011)). However, non-specific adsorption / desorption kinetics of oxalate species on $nTiO_2$ at pH 4 (Young and Mcquillan 2009) reveals a complex adsorption mechanism. Thus, a likely complex partitioning of HEPES buffer with oxalic acid limits adequate explanation of the UQ trend at this level.

h) Charged Ions

Figure 5.11 shows estimated increase in adsorbates of SRHA > SRFA > NOM > MES adipic acid and a decrease in HEPES oxalic acid in the presence of monoatomic and mixed ions.

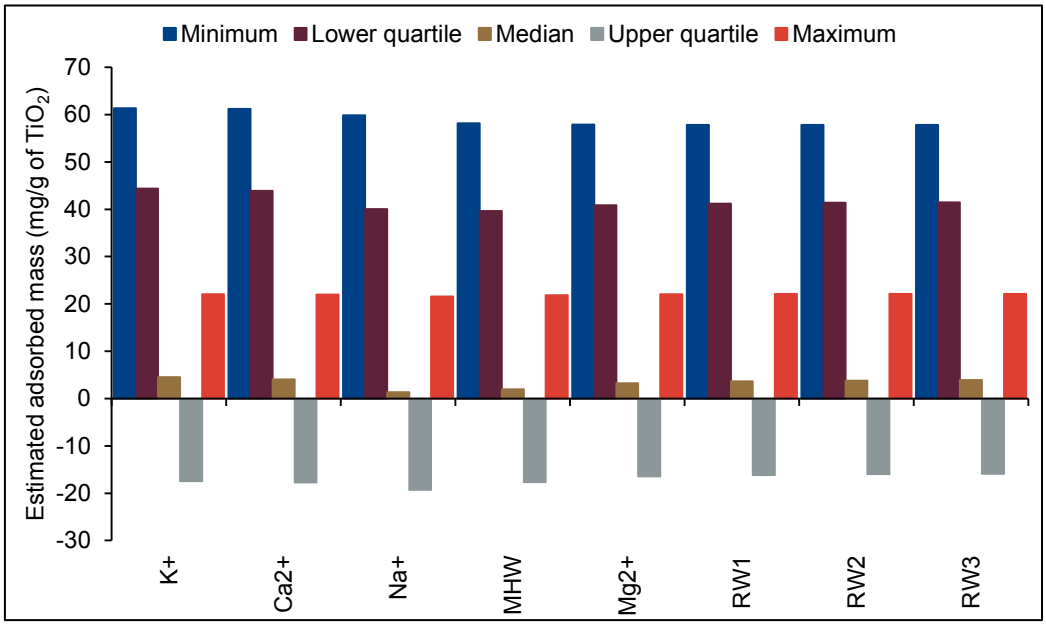


Figure 5.11 Profiles of adsorbed mass estimated by varying charged ions.

There is minimal individual contribution variation in projected adsorbates but the sequence is similar to that listed in Table 5.4 using the UQ and maximum data values. However, the trends simulated using the minimum, LQ and median data are non-specific except for the leading contribution of K^+ > Ca^{2+} with a ranking dissimilar to the one listed in Table 5.4. However, simulations suggest potential application of free and complexed ions in predicting adsorption behaviour of ENMs in diverse aquatic environment.

i) Ionic strength

There is lack of a distinct trend in the relative importance of SOM and NOM on adsorption. However, an increase in ionic strength marginally increases positive adsorbates of SRHA >

SRFA > NOM > MES adipic acid, and negative adsorbates of HEPES oxalic acid (Fig. 5.12), which agrees with published findings (Site 2001) attributable to ‘salting out’ effects.

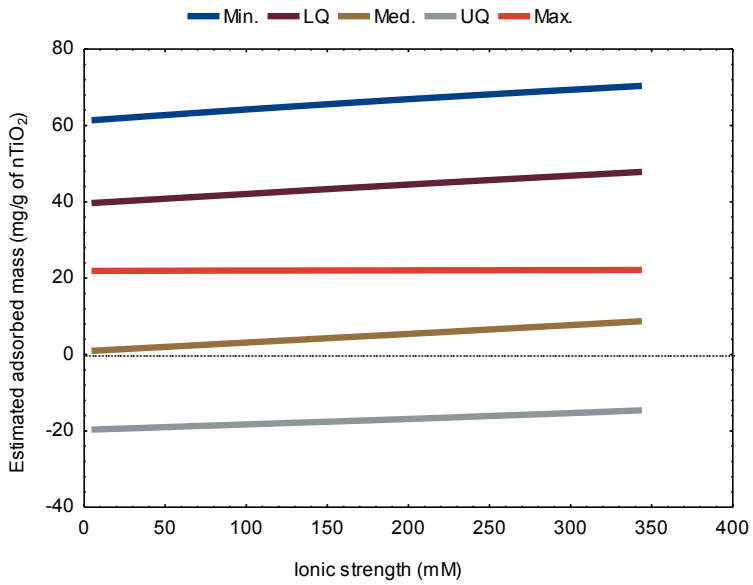


Figure 5.12 Profiles of adsorbed mass estimated by varying the ionic strength.

While the SRHA and SRFA patterns correspond to findings (Chen et al. 2012; Erhayem and Sohn 2014b), the NOM pattern contrasts findings by Chekli et al. (2015) and Chekli et al. (2015) because of low NOM and high Ca^{2+} concentrations. Thus, the conditions simulated by maximum data values play a role in the *ca.* constant increase in NOM adsorbates. In contrast, adsorption of oxalic and adipic acid on $nTiO_2$ determined at constant ionic strength (Hug and Bahnemann 2006; Pettibone et al. 2008; Young and Mcquillan 2009) limits inference of the findings to the extracted trends, thus, the need for further research.

5.4 Significance of the Model to Aquatic Environment

The model findings confirm that in an aquatic environment, various adsorptive including buffers will compete to attach on the $nTiO_2$ and in the end influence the surface characteristics of the ENM. Thus, the MLP neural networks model is applicable in situations having diverse competing factors that influence adsorption. Moreover, the extracted contribution of mixed oxides of $nTiO_2$ to predicting adsorption shows model potential in approximating similar phenomenon on diverse types and nanostructures of ENMs released into aquatic systems.

The application of hypothetical values has shown the capability of the model to extract linear and non-linear patterns that simulate the effect of each input variable in predicting the mass

of organic matter adsorbed on $nTiO_2$ using both qualitative and quantitative data. Thus, in the absence of $nTiO_2$-adsorptive-water partitioning constants, the MLP-NN can estimate adsorption behaviour using qualitative variables. The MLP model developed in this study is inductive, applicable in estimating similar associations for other ENMs and coupled with dynamic models; explain adsorption-influenced mobility, bioavailability, and toxicity of ENMs in the aquatic environment.

Studies have shown that electrostatic energy promotes the interaction between ENMs coated with organic matter and organisms leading to enhanced dispersion in an aqueous environment (Battin et al. 2009; Horst et al. 2010; Chowdhury et al. 2012). Thus, the potential predict adsorbates on ENMs can explain the likely long likely persistence and long-term effects in the aquatic environment, thus, the MLP-NN is usable tool that can support decision making on risk management of ENMs in aquatic environment.

5.5 Conclusion

An ensemble neural networks model developed using three physicochemical properties of $nTiO_2$ and six abiotic factors adequately predicted the quantity of ten types of organic matter adsorbed on the ENM. The relative importance of the input variables to the response hinges the fitted model to fundamental scientific principles of adsorption kinetics. The high ranking of temperature > duration of exposure > types of organic matter suggest thermal properties of aqueous environment and transfer kinetics determine the strength of bonds between organic matter and ENMs. Moreover, the location of atoms in the crystal structure of $nTiO_2$ facilitates adsorption bonding at the reactive sites more than the primary size, morphology, and concentration of the ENM. The least variables contributing to adsorption were pH > charged ions > ionic strength, implying certain thresholds of exposure conditions control adsorption rates and amounts of adsorbates.

The findings show that simulated exposures inadequately explain likely adsorption kinetics in the aquatic environment to support research in predictive modelling. Moreover, higher influence of K^+ on adsorption than Ca^{2+} contrasts published literature, perhaps attributable to cation exchange capability with different adsorptive that affects binding affinities. The model consistently depicted the least ranked variables and interchanged association of the first six variables after varying the perturbation value. Simulated input versus output relationship trends reveal competency of neural networks to forecast adsorbed mass of various types of organic matter on $nTiO_2$ under different exposure conditions.

Chapter 6 - Prediction of Hydrodynamic Size of $nTiO_2$ in Water

6.1 Summary

This Chapter presents findings addressing the second part of the second objective and the first part of the third objective by first approximating the hydrodynamic size of $nTiO_2$ particles using ten input variables, and second, simulating the relationship between adsorption and aggregation. The objectives provide answers to the second part of the research question that sought to establish associations between properties of engineered nanomaterials (ENMs) and the environment, as causal factors, and aggregation.

For the first time, application of multilayer perceptron neural networks model predicted size evolution of $nTiO_2$, and simulated likely relationships between input variables and the response, in an aqueous environment. The input variables comprise composition of anatase crystals, surface area concentration (SAC), and surface coating as properties of $nTiO_2$; and pH, ionic strength, temperature, duration of exposure, the types and initial concentration of organic matter, and charged ions as environmental abiotic factors.

The average training, cross-validation, and out-of-sample testing set mean square error (MSE) of standalone models was 0.174 ± 0.015, 0.210 ± 0.030, and 0.397 ± 0.042, respectively. An ensemble model selected based on validation set error predicted the hydrodynamic size with an MSE of 0.266, which improved performance of standalone models by 32.8%. The prediction $R^2 = 0.732$ depicts good model generalization and approximation of new experimental data.

The surface chemistry of ENMs and abiotic factors affect hydrodynamic size more than size, concentration, morphology, and crystal structure of $nTiO_2$, and initial concentration of organic matter. The sequence of significant variables is surface coating > temperature > ionic strength > types of organic matter > duration of exposure > pH > charged ions > SAC > initial concentration of organic matter > composition of anatase crystals.

The effect of diverse ions and ligands on $nTiO_2$ surface charges result in complex aggregation behaviour and size evolution in an aqueous environment. Contrary to published literature, an increase in temperature also decreases aggregation based on exposure conditions. A nonlinear dependence of hydrodynamic size on organic adsorbates suggests a

broad influence of the latter on size evolution, dispersion, and settlement of ENMs. Results offer insights to support future development of dynamic models on aggregation behaviour.

6.2 Introduction

In an aqueous environment, favourable conditions promote particle-particle and particle-solid interactions resulting in mono- and hetero-aggregation, respectively. In the presence of organic ligands, hetero-aggregation and generation of ENMs-ligand fractals would be prevalent, which is the subject of discussion in this Chapter. The minimum content of organic matter is 0.0135 mg/ℓ in purified aqueous media (refer to Appendix 4.3). Chapter 4 discusses in depth properties of $nTiO_2$ and abiotic factors investigated as input variables in this study.

The atom configurations and chemical bonding underscore the relevance of anatase crystal structure on aggregation. The SAC is a primary physical property and the basis for size evolution of ENMs. The pH and ionic strength determine interfacial reactivity and transformation of ENMs' sizes. An increase in temperature would decrease the viscosity of aqueous media, which reduces the magnitude of zeta potential, and vice versa (refer to Eqn. 4.3). The temperature effects on aggregation of $nTiO_2$, conditions investigated in this study lie between 19.85 and 25.00°C, inclusive. Additionally, the concentration of organic ligands and duration of exposure determine the quality of aggregates formed.

The primary capping of ENMs before exposure represents the surface chemistry and potential reactivity. Under different conditions, coated and uncoated (bare) ENMs acquire diverse magnitudes of surface energy, an inconsistent characteristic even for particles from the same batch. In this study, inadequate data hinders application of this feature assumed constant by limiting effects on the evolution of size to surface capping agents. Nominal levels of this variable and % of training examples in the data subset include bare ENMs (98.6), and those coated with $Al(OH)_3$ and polydimethylsiloxane (PDMS) (AlPDMS) (0.7) and $Al(OH)_3$, PDMS, glycerin and other organics (AlPDMSGl) (0.7). Excluded are effects of impurities.

Studies have investigated the influence of various types and concentrations of definite ligands on aggregation of ENMs (Appendix 4.3). However, the organic, inorganic, and biological components of aquatic systems would uniquely interact with ENMs. Thus, this study investigates the effect of total carbon (TC) in the presence and absence of biological organism, natural organic matter (NOM), and Suwannee River fulvic and humic acids (SRFA and SRHA) (Table 4.7 in Chapter 4) to predicting hydrodynamic size. Whereas TC only excludes complex elements found in surrogate organic matter (SOM), namely, SRHA, SRFA

and NOM, the presence of biological organism adds diversity of a hydrophobic biochemical ligand.

The training data used in this study were generated using natural, standard synthetic and purified aqueous water. Section 4.5.2 and the previous Chapter describe charged ions as an input variable representing diverse properties of aqueous media. To enhance learning, 28 nominal levels of charged ions investigated in this study include the following;

1. Mixed ions in river water (5), lakewater (4), sea water (3), groundwater (2), brackish water (2), tap water (1), and peat bog water (1).
2. Mixed ions in mesocosm (1), standard toxicity test water (1), and standard synthetic water (4).
3. Four mono ions (K^+, Na^+, Ca^{2+}, and mixed Fe^{3+} and Na^+) in purified water.

The diverse locations categorize similar natural water systems having identified as well as presumed unidentified distinctive properties. The ordinal numbers (in brackets) signify splits that classify nominal levels. Based on packing density, relative spaces covered by irregular fractals determine the size of pervious or impervious hetero-aggregates formed. The hydrodynamic sizes were mainly measured using dynamic light scattering (DLS) technique at an instantaneous time up to 504-hours under fast and slow aggregation regimes. The DLS technique measures spherical particle sizes in fluids based on Stokes-Einstein equation (Eqn. 6.1).

$$d_H = \frac{k_B T}{\eta D} \tag{6.1}$$

Equation 6.1 correlates hydrodynamic size, (d_H), of a particle in Brownian motion directly to temperature (T) and Boltzmann constant (k_B), and indirectly to fluid viscosity (η) and diffusion coefficient of suspended particles (D). One study, (Chekli et al. 2015) combined DLS and flow-field flow fractionation (FIFFF) techniques to measure smaller and larger particle sizes during slow and fast aggregation, respectively. The FIFFF is a high resolution elution technique capable of determining a broad range of sizes ≤ 100 µm independent of particle density (Eqn. 6.2) (Shimpf et al. 2000).

$$d_H = \frac{2t_r k_B T v_o}{t_o \pi \eta w_c Q_o} \tag{6.2}$$

where (t_r), is residence time of particles retained in fluid, (v_o) is the void volume of micro-channel, (t_o) is the residence time of particles not retained, (w_o) is the flow micro-channel width and (Q_o) is the volumetric flow rate. Both techniques estimate hydrodynamic size based on particle velocities in fluid micro-channels that underscore size correlation with residence time. In DLS, small particles move faster than large ones, with former attaining lower decay curves. The FIFFF dispersive technique averages hydrodynamic size from clusters of distributed sizes. Non-spherical shapes or surface adsorbed layers account for larger sizes measured, which enhances polydispersity (Goodwin 2004).

Diverse properties of aqueous media and analytical techniques used generated response data that exhibits a wide range of average hydrodynamic sizes of ENMs. The distribution of different sizes measured in related media displays large variances (Table 4.7 in Chapter 4). Models fitted using unbalanced data are inclined to utilize outliers with large deviations from the sample mean, which reduces accuracy. In this study, implementing stratified sampling aimed at regulating and evenly distributing different magnitudes of the response data.

The fitting of the model trained in Chapter 5 to input data used in this study, except surface coating, tested the competency in estimating mass coverage of organic matter on ENMs to establish a relationship between the two kinetic responses. Applications of study findings are similar to those outlined in the previous Chapter, which relate aggregation of $nTiO_2$ to mobility, bioavailability, toxicity, and management of ENMs in aquatic environment.

6.3 Results and Discussion

The methods presented in Section 3.2, and procedures described in Chapter 5, apply to prepare, and partition training data; determine suitable weight decay value and network configurations used to train models; select ensemble model to implement, and evaluate variable importance in predicting hydrodynamic size. Similarly, the tanh and identity were transfer functions used in the hidden and output layers, respectively. The number of iterations was set at 1,000; hidden neurons set from 4 to 9, inclusive, totaling 6 different network configurations; and 5 models retained after training the networks.

Prior pre-partitioning using a 4:2:6:1 sampling ratio distributed the response data with large variances (refer to Table 4.7) into comparable stratum holding 130-500, 501-1,000; 1,001-10,000; and 10,001-28,000 nm size ranges, respectively (Table 6.1). The stratum ratio is the number of training examples in each sub-group (count) divided by the minimum number of training examples, in this case, 23 values from the 10,001-28,000 nm range (Table 6.1).

Table 6.1 Distribution of training examples into comparable stratum

Experimental hydrodynamic size (nm)						
Count	Minimum	Maximum	Median	Mean	Std. Dev.	Stratum ratio
81	131.80	500.00	311.11	300.96 ± 24.65	111.50	4
45	510.00	1,000.00	761.11	766.35 ± 45.40	151.10	2
141	1,007.50	8,713.00	1,529.41	1,782.93 ± 163.81	983.88	6
23	10,256.00	27,665.90	20,565.00	20,038.06 ± 1,966.10	4,546.60	1

Ten percent of the data were set aside for out-of-sample testing of trained models, the remainder (90.0%) apportioned into training (209), and cross-validation (52) sets using 1, 562, 1,564, 5,798, and 8,764 random subsampling seeds during parallel model building. Preliminary model training at fixed random seed of 1,564 and 7 hidden neurons, and varied weight decay (α) values between 10^{-7} and 10^{-1} in the hidden layer and 10^{-8} and 10^{-2} in the output layer, determined 10^{-4} and 10^{-5} as α values for use in the input and output layers, respectively (Table 6.2). The models trained using the selected α values exhibited the minimum validation set error.

Table 6.2 Weight regularization values for hydrodynamic size-based models

Weight decay (α)		Parallel training MSE		Ensemble
Hidden layer	Output layer	Validation	Testing	Testing
10^{-7}	10^{-8}	0.519 ± 1.057	2.260 ± 5.094	0.594
10^{-6}	10^{-7}	0.191 ± 0.395	0.919 ± 1.330	0.266
10^{-5}	10^{-6}	0.218 ± 0.037	0.635 ± 0.360	0.336
10^{-4}	10^{-5}	0.207 ± 0.082	0.600 ± 0.307	0.230
10^{-3}	10^{-4}	0.240 ± 0.070	0.544 ± 0.280	0.487
10^{-2}	10^{-3}	0.426 ± 0.233	0.839 ± 0.056	0.832
10^{-1}	10^{-2}	0.465 ± 0.280	0.948 ± 0.010	0.946

Table 6.3 summarizes the performance of 25 models retained for each of the 6 MLP-NN topologies evaluated. Figures 6.1 and 6.2 depict the variances of validation and testing set error of the MLP-NN models trained using different initial random seeds for each topology. Similar to observations made in the previous Chapter, models that display the minimum

average validation set error do not essentially yield the best out-of-sample prediction (Table 6.3). However, the results show that combining models improves their prediction capability that varies with the network configuration used in training the models. Thus, fitting new data for similar input variables in the ensemble models listed in Table 6.3 would sufficiently predict hydrodynamic size. The MLP-NN structure having 4 hidden neurons displays the minimum average validation set error and a narrow spread of error variances except for few outliers (Figures 6.1 and 6.2). This model was henceforth, selected for implementation and evaluation of sensitivity analysis discussed in subsequent sections.

Table 6.3 Summary of standalone and ensemble model performance

	Standalone model MSE (0.95 confidence interval)			Ensemble model MSE		Ensemble
Topology	Training	Validation	Testing	Testing	Reduction (%)	R^2
MLP 10-4-1	0.174 ± 0.015	0.210 ± 0.030	0.397 ± 0.042	0.266	32.8	0.732
MLP 10-5-1	0.170 ± 0.014	0.225 ± 0.034	0.397 ± 0.046	0.240	39.5	0.755
MLP 10-6-1	0.157 ± 0.014	0.222 ± 0.037	0.406 ± 0.083	0.212	47.9	0.780
MLP 10-7-1	0.160 ± 0.011	0.220 ± 0.034	0.372 ± 0.053	0.217	41.6	0.780
MLP 10-8-1	0.163 ± 0.017	0.223 ± 0.032	0.449 ± 0.092	0.220	51.0	0.770
MLP 10-9-1	0.151 ± 0.010	0.221 ± 0.035	0.357 ± 0.136	0.171	52.1	0.821

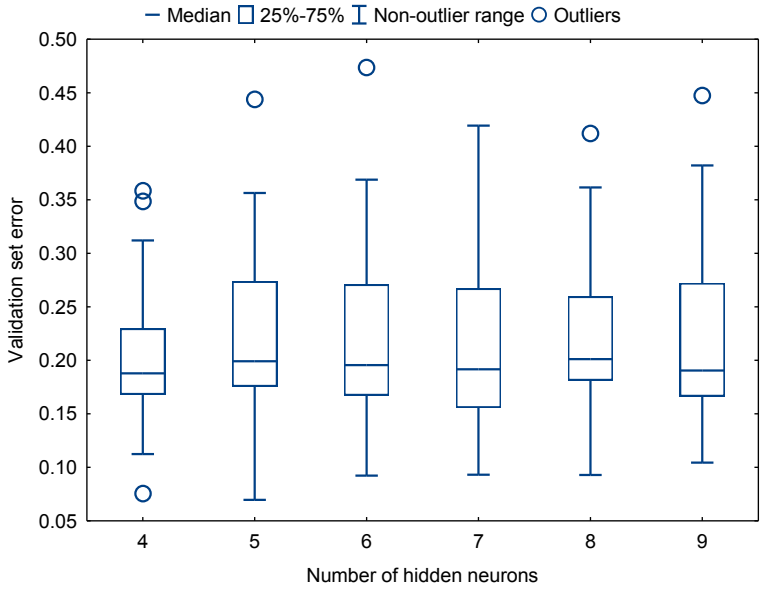


Figure 6.1 Validation set errors of networks having a different number of hidden neurons.

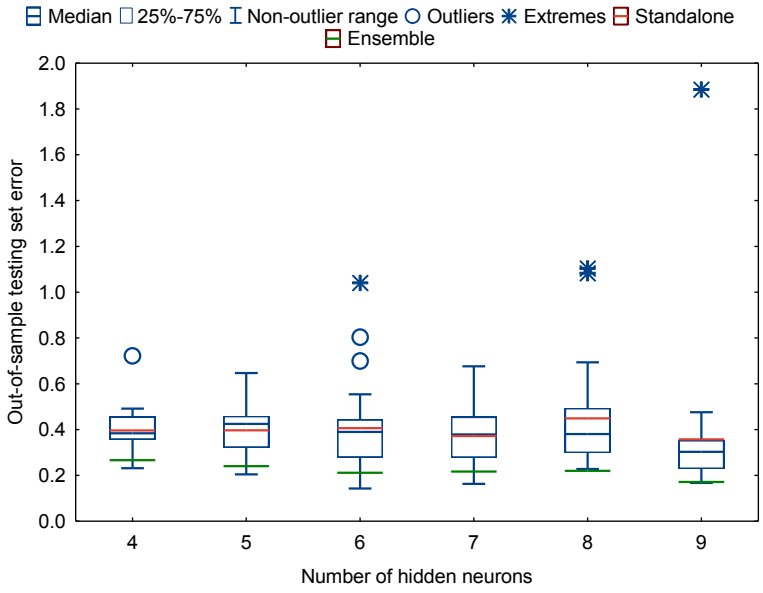


Figure 6.2 Testing set errors of networks having a different number of hidden neurons.

A scatter comparison between predicted and experimental data shows adequate prediction of hydrodynamic size of $nTiO_2$ with $R^2 = 0.732$ (0.95 prediction interval) (Fig. 6.3) indicating good approximation of new experimental data by the ensemble model.

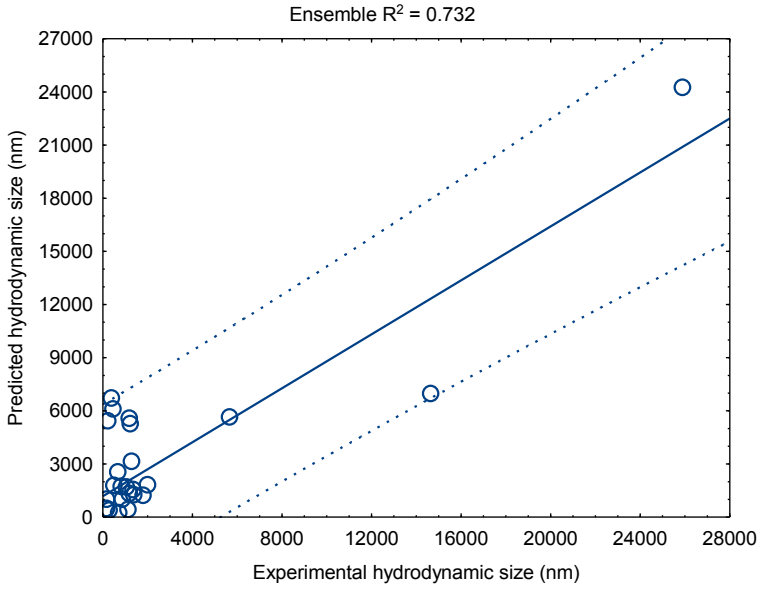


Figure 6.3 Ensemble prediction of the $nTiO_2$ hydrodynamic size in aqueous environment.

A residual analysis demostrates independence $(R^2 = 0.004)$ of predicted versus residual hydrodynamic size (Fig. 6.4) (*Top*). A left-skewed residual values' distribution suggests higher over-prediction (*ca.* 62.1%) than under-prediction of the response (Fig. 6.4) (*Bottom*).

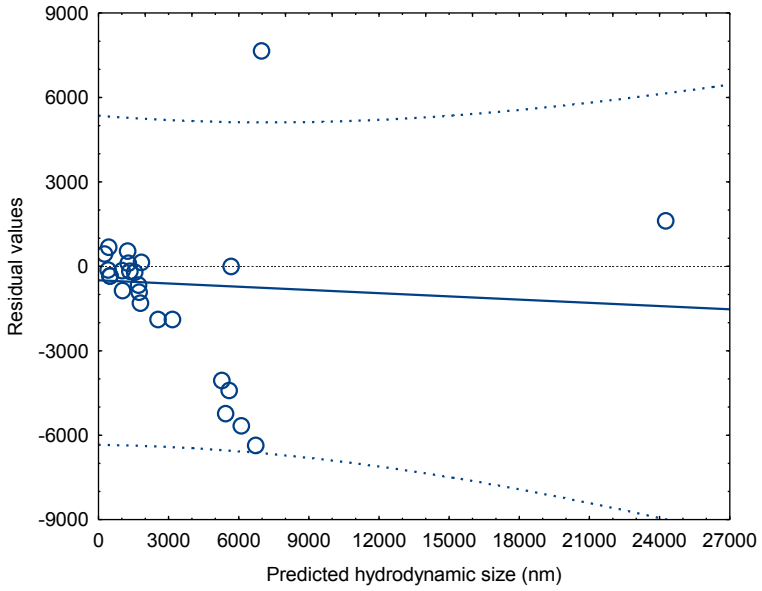
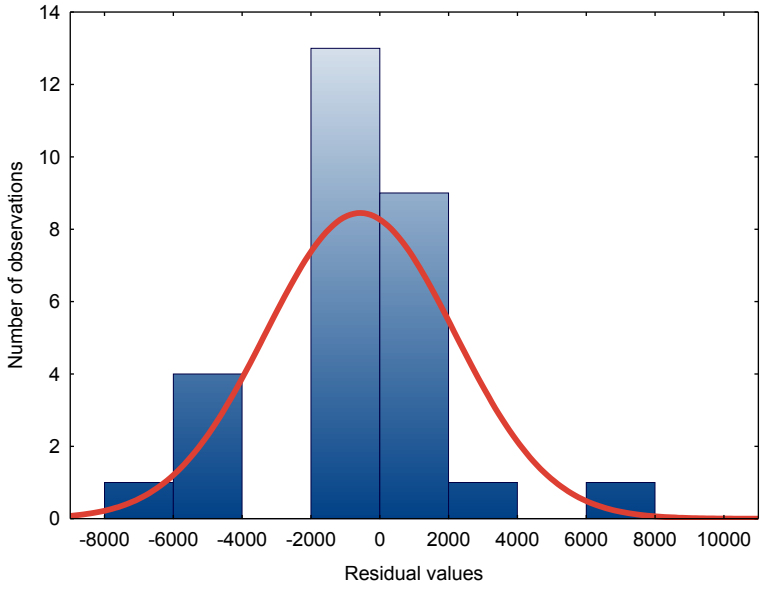


Figure 6.4 Predicted versus residual values (*Top*) and normality of residual values (*Bottom*).



6.3.1 Relative Importance of Input Variables

The implemented ensemble model had a mean square error (MSE) of 0.317 based on the training and validation sets. The method by Francis (2001) (refer to Section 3.2.4 in Chapter 3 and Section 5.3.2 in the previous Chapter) was used to determine the relative importance of input variables in predicting hydrodynamic size. Table 6.4 presents the MSE of new models computed by perturbing the input variables in the fitted ensemble model. The relative importance ranks show a 7.02- and 1.01-fold increment in MSE of the full model by

perturbing the most (surface coating) and least (composition of anatase crystal) influential input variable after perturbation, respectively (Table 6.4).

Table 6.4 Relative importance of input variables in predicting hydrodynamic size of $nTiO_2$

Variables	MSE	Sensitivity (%)	Rank
Surface coating	2.228	85.8	1
Temperature (°C)	1.586	80.0	2
Ionic strength (mM)	0.718	55.9	3
Types of organic matter	0.472	32.9	4
Duration of exposure (hours)	0.420	24.6	5
pH	0.416	23.8	6
Charged ions	0.401	20.9	7
Surface area concentration (SAC) (m^2/ℓ)	0.344	7.9	8
Initial concentration of organic matter (mg/ℓ)	0.330	3.9	9
Composition of anatase crystals (%)	0.321	1.3	10

The surface coating (or the surface chemistry of $nTiO_2$) and abiotic factors, except for initial concentration of organic matter, have higher influence on aggregation than the size, concentration, and crystal structure of the ENM. Hence, particle-liquid interface reactions likely play a key role in the evolution of nanoparticle sizes. Moreover, concentration of organic matter stabilizes or destabilizes nanoparticles depending on the exposure conditions, which explains low sensitivity.

Individually, the sensitivity of $nTiO_2$ coated with AlPDMS and AlPDMSGI is *ca.* 300-fold higher than that of uncoated $nTiO_2$ while the sensitivity of surrogate organic matter (SOM) is greater than that of NOM to predicting hydrodynamic size (Table 6.5). A ligand association between degraded capping agents and reactive $nTiO_2$ surface (Auffan et al. 2010) can create large sized asymmetrical assemblages.

It is important to point out that 98.5% of training examples comprise values of bare $nTiO_2$ mostly used in experimental studies. Therefore, uncertainty in extracting individual sensitivities after substituting uncoated with coated values of $nTiO_2$ in the fitted ensemble model is likely. However, the results suggest potential underestimation of size evolution of ENMs released from nanoproducts into the aquatic environment, most of which are coated, by the use of uncoated nanoparticles in aggregation studies. In contrast, the use of SOM

and salts in aggregation studies may over-estimate hydrodynamic size of $nTiO_2$ in the presence of NOM.

Table 6.5 Relative importance of individual surface coating and types of organic matter in predicting hydrodynamic size of $nTiO_2$

Input variable	MSE	Sensitivity (%)	Rank
Surface coating material			
AlPDMSGl	3.210	90.1	1
AlPDMS	3.160	90.0	2
Bare (uncoated surface)	0.320	0.3	3
Types of organic matter			
SRHA	0.596	46.8	1
Total carbon and organism	0.508	37.5	2
SRFA	0.506	37.3	3
Total carbon (TC)	0.425	25.4	4
Natural organic matter (NOM)	0.328	3.4	5

The monatomic ions exhibit higher influence in the sequence $K^+ > Ca^{2+} > Na^+$ than mixed ions to predicting hydrodynamic size (Table 6.6). The remainder of assorted mixed ions displays a 0.04 - 2.3% change in sensitivity from preceding values, implying minimal variation in their influence on aggregation. The lack of an influential trend in standard synthetic and natural water suggest likely simultaneous effects of organic matter and mixed ions in balancing surface charges of $nTiO_2$ that uniquely hinder aggregation.

Considering the large number of nominal levels, two side-by-side contrasts of applied numerical codes and sensitivity outcomes established a 17.8% likely ordering. The evaluations involved comparing smallest-to-largest quantitative scores of nominal levels versus, (i) smallest-to-largest MSE values; and (ii) largest-to-smallest MSE values. The ordered sensitivities could be coincidental, but suggest a procedure in data preparation that may contribute to uncertainties of neural networks predictions using categorical variables, which performs best by learning from numerical data.

Table 6.6 Relative importance of individual charged ions in predicting hydrodynamic size of $nTiO_2$

Charged ions	MSE	Sensitivity	Rank	Charged ions	MSE	Sensitivity	Rank
K^+	0.497	36.2	1	$TP_{(Mixed)}$	0.394	19.5	15
Ca^{2+}	0.494	35.8	2	$LW_{1(Mixed)}$	0.393	19.4	16
Na^+	0.458	30.8	3	$VSW_{(Mixed)}$	0.393	19.4	16
$RW_{4(Mixed)}$	0.407	22.2	4	$SW_{1(Mixed)}$	0.393	19.3	18
$GW_{(Mixed)}$	0.405	21.8	5	$VHW_{(Mixed)}$	0.393	19.2	19
$MHW_{(Mixed)}$	0.405	21.8	5	$SW_{(Mixed)}$	0.388	18.2	20
$LW_{(Mixed)}$	0.403	21.3	7	Fe^{3+} and $Na^+{}_{(Mixed)}$	0.386	17.8	21
$RW_{5(Mixed)}$	0.401	20.9	8	$RW_{1(Mixed)}$	0.385	17.8	21
$LaBW_{(Mixed)}$	0.399	20.6	9	$RW_{2(Mixed)}$	0.384	17.4	23
$TvBW_{(Mixed)}$	0.398	20.4	10	$RW_{3(Mixed)}$	0.383	17.3	24
$SimF_{(Mixed)}$	0.397	20.1	11	$HW_{(Mixed)}$	0.380	16.6	25
$IsoLeF_{(Mixed)}$	0.395	19.8	12	Algal $TM_{(Mixed)}$	0.371	14.6	26
$PBW_{(Mixed)}$	0.394	19.6	13	$MesoF_{(Mixed)}$	0.366	13.4	27
$GW_{1(Mixed)}$	0.394	19.6	13	$SB\_SW_{(Mixed)}$	0.364	12.9	28

Key: MHW, synthetic moderately hard water; SW, SB_SW and SW_1, seawater from different regions; $RW_{1\text{-}5}$, river water from different regions; Algal TM, standard algal test media; HW, synthetic hard water; VHW, synthetic very hard water; VSW, synthetic very soft water; LW, LW_1, IsoLeF and SimF, lake water from different regions; TW, tap water; GW and GW_1, ground water from different regions; PBW, peat bog water; and TvBW and LaBW, brackish water from different regions.

6.3.2 Simulated Explanatory Input-Hydrodynamic Size Relationships

Table 6.7 presents hypothetical data values used to estimate the model sensitivity to changes in input variables according to Lek et al. (1996) technique (refer to Section 3.2.4 of Chapter 3 and Sections 5.3.3 in the previous Chapter). The minimum data values simulate a 1:1 electrolyte acidic condition favouring steric stabilization (French et al. 2009). Conversely, the maximum data values simulate an ambiguous exposure of capped $nTiO_2$ in an environment having multiple interacting factors, including mixed ions, biological organism, ionic strength of seawater, and high concentration of organic matter, which can uniquely control surface charge and stability of the ENM. The data values for temperature, types of organic matter, surface coating, and composition of anatase crystal exhibits repetitions.

Table 6.7 Hypothetical input data used to predict hydrodynamic size of $nTiO_2$

Input variable	Minimum	LQ	Median	UQ	Maximum
Surface area concentration (m^2/ℓ)	0.01	0.52	2.00	2.58	76.65
Temperature (°C)	19.85	22.00	23.5	25.00	25.00
Concentration of organic matter (mg/ℓ)	0.01	1.16	1.85	6.57	114.21
Duration of exposure (hours)	0	0.43	1	24	504
Types of organic matter	SRHA	NOM	NOM	NOM	TC and organism
Charged ions	K^+	Ca^{2+}	SB-SW	Fe^{3+} and Na^+	RW_4
Surface coating	Bare	Bare	Bare	Bare	AlPDMSGl
Composition of anatase crystals (%)	0.0	80.0	82.0	82.0	100.0
pH	4.00	6.53	7.60	8.05	8.38
Ionic strength (mM)	0.30	2.50	5.15	707.00	750.00

LQ, lower quartile; UQ, upper quartile; RW_4, river water in a specific site (RW in (Chekli et al. 2015).

The subsequent Sections (a) to (j) present parallel discussion of findings and synergy from the two sensitivity analyses with references made to Tables 6.4 to 6.6. The positive and negative output values, the latter mostly estimated using the minimum and maximum data, indicate an increase and decrease in hydrodynamic size with respect to the input variables, respectively. Moreover, since the data do not represent definite synthetic or natural exposure conditions reported in published literature, the predicted response output trends are linked to current knowledge about aggregation of nanoparticles where applicable.

a) Surface Coating

There is a similarity in the sequence of relative importance of capping agents presented in Section 6.3.1 and Figure 6.5. The capping agents enhance stability of ENMs (Cao 2004; Hornyak et al. 2009). Nevertheless, the model estimates higher size evolution of capped than uncapped $nTiO_2$ in aqueous environment. The degradation of AlPDMS and transformation of surface properties of $nTiO_2$ (Auffan et al. 2010) affect stability of the ENM in the presence of organic matter (Labille et al. 2010). Thus, the extracted relationships demonstrate the fitness of characterizing properties of capping agents on ENMs and applying them as input variables in predictive modelling to approximate size evolution.

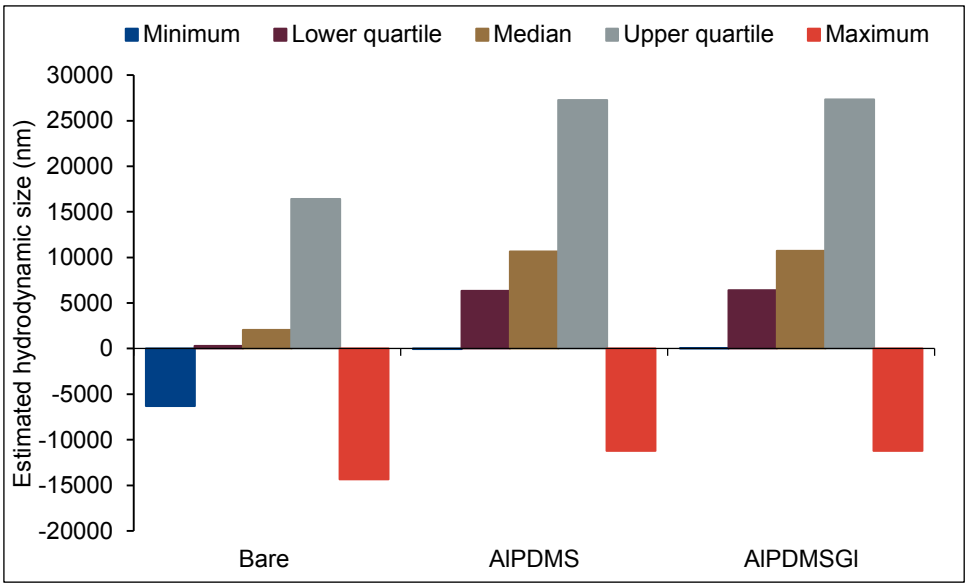


Figure 6.5 Contribution of surface coating material to predicting the hydrodynamic size of $nTiO_2$.

b) Temperature

The LQ-UQ data simulate an increase in hydrodynamic size with increasing temperature, with the UQ trend showing negative and positive response outputs (Fig. 6.6).

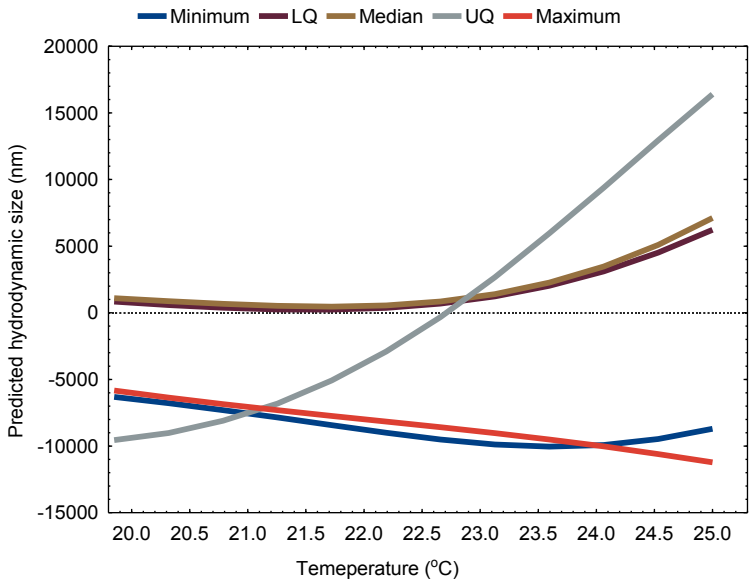


Figure 6.6 Contribution of temperature to predicting hydrodynamic size.

In monatomic low NOM exposure conditions, an increase in temperature lowers the magnitude of zeta potential (Chen et al. 2012) and enhances thermal disruption of weak aggregate bonds (Zhou et al. 2012) leading to aggregation and size increment. Conversely,

the decrease in hydrodynamic size with increasing temperature (minimum and maximum trends) is uncommon and needs further investigation. Thus, the model has revealed possibilities of seasonal temperature variations affecting aggregation, size evolution and partitioning of ENMs in an aquatic environment.

c) Ionic strength

The hydrodynamic size of $nTiO_2$ increases (LQ, median and UQ trends) or decreases (minimum and maximum trends), with increasing ionic strength (Fig. 6.7). The LQ-UQ trends are in agreement with published findings (Keller et al. 2010; Ottofuelling et al. 2011; Chekli et al. 2015) influenced by electrostatic attraction because of compressed electrical double layer and decrease in magnitude of zeta potential (Thio et al. 2011). The minimum and maximum trends simulate conditions that promote charge neutralization and bridging effects, respectively, that decrease aggregation (Zhu et al. 2014).

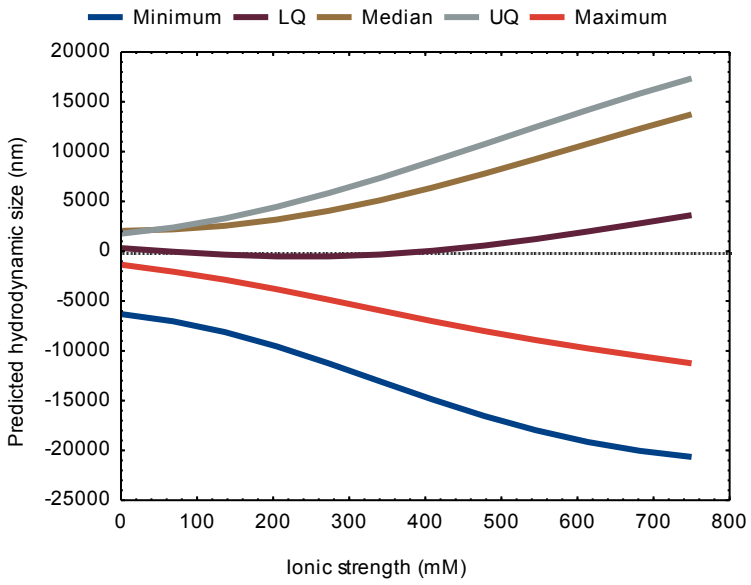


Figure 6.7 Contribution of ionic strength to predicting hydrodynamic size.

d) Types of Organic Matter

The sequence of ranking types of organic matter presented in Table 6.5 is different from that simulated in Figure 6.8. The order TC + organism > TC > NOM > SRFA > SRHA is similar under the median, UQ, and maximum conditions, which shows higher stability in the presence of SOM than NOM contrary to findings (Ottofuelling et al. 2011). Further, the order SRHA > SRFA > NOM > TC > TC + organism, and SRHA > NOM > TC > SRFA > TC +

organism, under the minimum and LQ conditions, respectively, suggest stability in the presence of organism that agrees with findings (Horst et al. 2010).

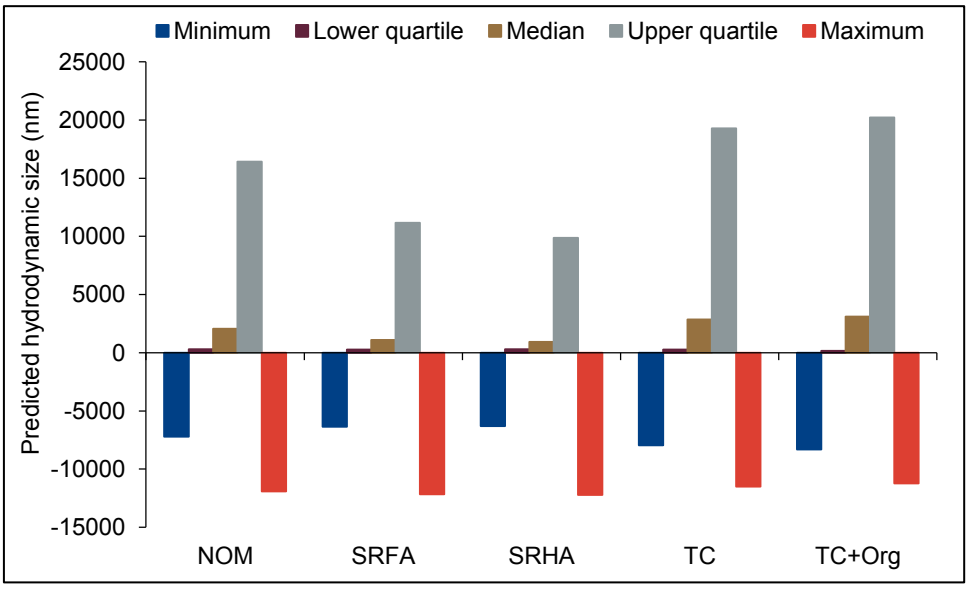


Figure 6.8 Contribution of the types of organic matter to predicting hydrodynamic size.

The projected sizes illustrate the intricate interactions between ENMs with organic and biological ligands contingent on properties of an aqueous environment, which result in stabilizing or destabilizing effects. The diverse properties and binding affinities of ligands described in the previous Chapter, and the intricacy of aggregate solids formed explain the predicted $nTiO_2$ sizes. Thus, qualitative properties of diverse aquatic ligands are promising predictors of ENMs' size evolution in the absence of known attachment efficiencies.

e) Duration of Exposure

The positively or negatively low hydrodynamic size at instantaneous exposure periods and decrease with increasing duration of exposure (Fig. 6.9) contrasts published literature (Zhang et al. 2008; Keller et al. 2010). The dataset consists response values measured under both diffusion-limited (≥ 10,000 nm after short exposure periods) and reaction-limited (≤ 10,000 nm after long exposure periods) aggregation regimes. Thus, based on the data used, the MLP-NN has ascertained fast aggregation at instantaneous exposure. The likely formed partitioning of the sizes would be (I) settlement of large solids, (II) dispersion of small sized solids that coalesce and increase in size to produce compact fractals, and (III) stabilized solids whose size does not significantly increase with time.

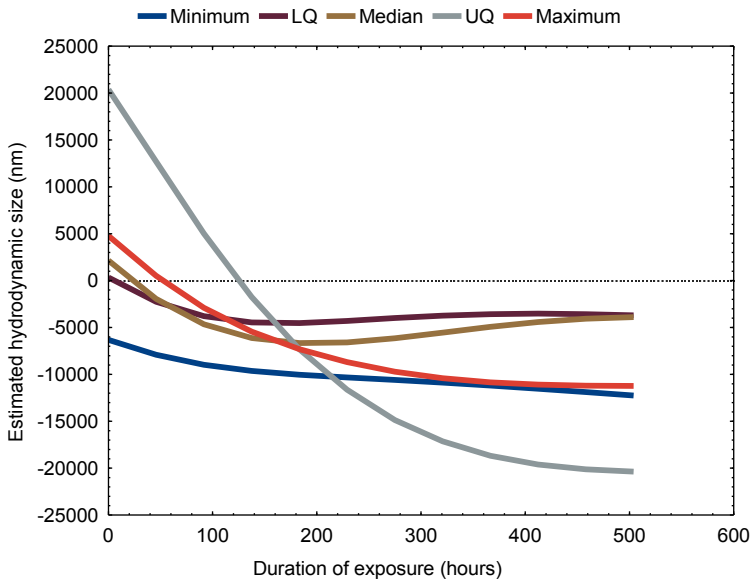


Figure 6.9 Contribution of the duration of exposure to predicting hydrodynamic size.

f) pH

As the pH increases, the decrease in hydrodynamic size increases in the positive direction and increases in the negative direction for the minimum and maximum trends, respectively (Fig. 6.10). The likely negative surface charge and steric stabilization of ENMs (Baalousha 2009) explain the maximum trend. In contrast, the size increases to a maximal peak and drops in the LQ-UQ trends, agreeing with published literature (French et al. 2009).

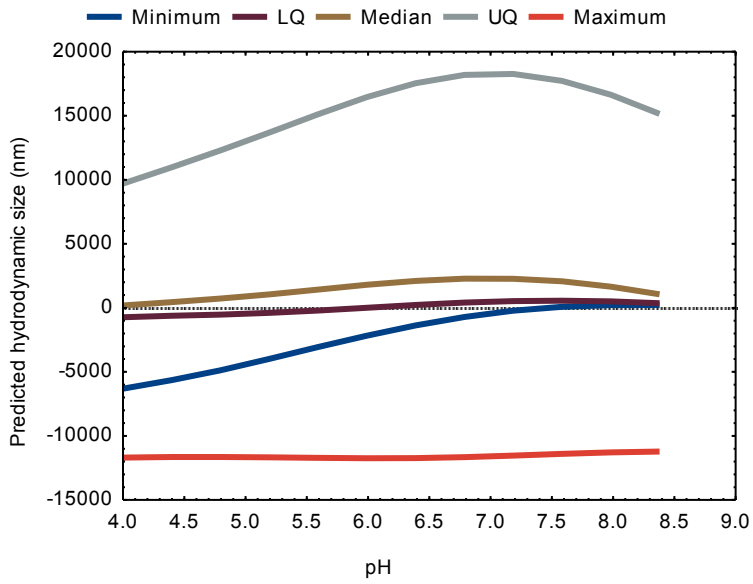


Figure 6.10 Contribution of the pH to predicting hydrodynamic size.

g) Charged ions

The minimum and maximum trends (Fig. 6.11) display model sensitivities that invert the sequence presented in Table 6.6. Conversely, the sequences of 5-most influential ions shown by the LQ, median and UQ is K^+ > Ca^{2+} > RW_4 > GW > LW, HW > RW_3 > RW_2 > RW_1 > SW, and MHW > Na_+ > SW > RW_1 > RW_2, respectively. The extracted trends demonstrate the complexity of measuring hydrodynamic size in diverse exposure systems.

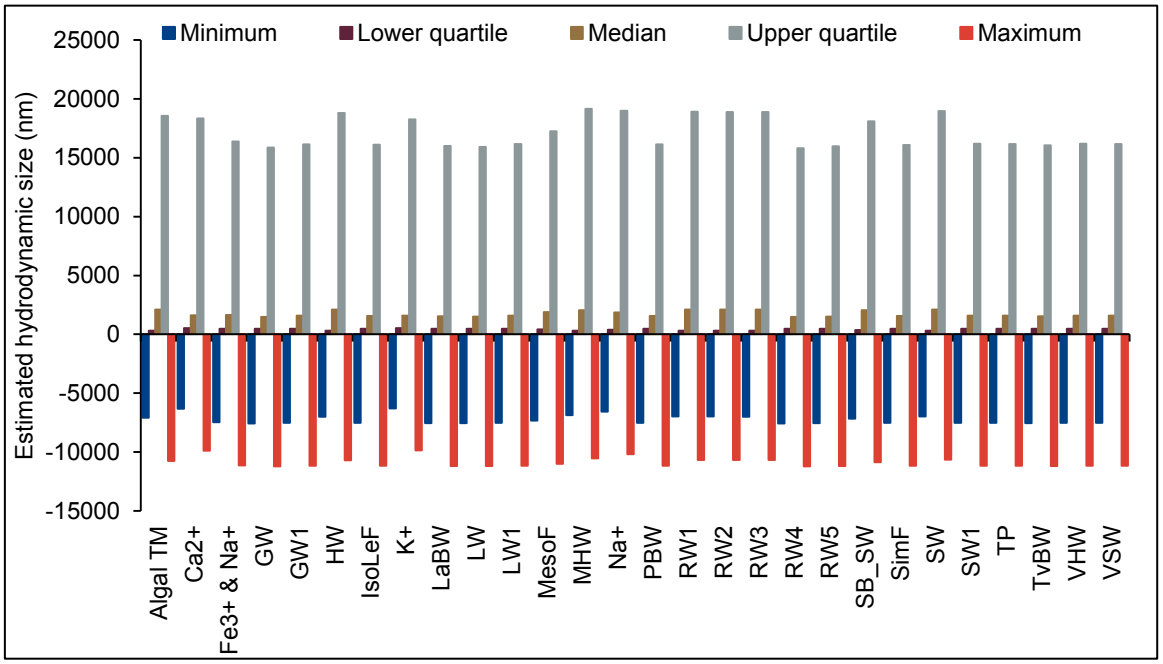


Figure 6.11 Contribution of charged ions to predicting hydrodynamic size.

The published literature associates the effects of ions on aggregation of ENMs based on measurements of single ions in aqueous suspensions. The hydrodynamic size trends have been reported in natural water (Keller et al. 2010; Ottofuelling et al. 2011; Sillanpää et al. 2011) but the synergy between various species of mixed ions on aggregation is undocumented. Moreover, the emphasis that is placed on the effects of Ca^{2+} (Chekli et al. 2015) omits the contribution of other ion species on aggregation in natural water. However, the MLP-NN model has shown the capability of utilizing diverse ions and extracting their individual contribution to predicting aggregation, which is an exercise that would be challenging to measure empirically.

h) Surface Area Concentration

The influence of SAC on aggregation is liable to simulated data values used to predict hydrodynamic size (Fig. 6.12). The size increases with an increase in SA in the negative

direction of the minimum and maximum trends and positive direction of the LQ trend. Conversely, the median trend displays a positive drop in size to a minimal peak and a rise in the size with increasing SAC. The UQ trend shows a drop in large aggregates of $nTiO_2$ formed at low SAC and subsequent decrease with increasing SAC.

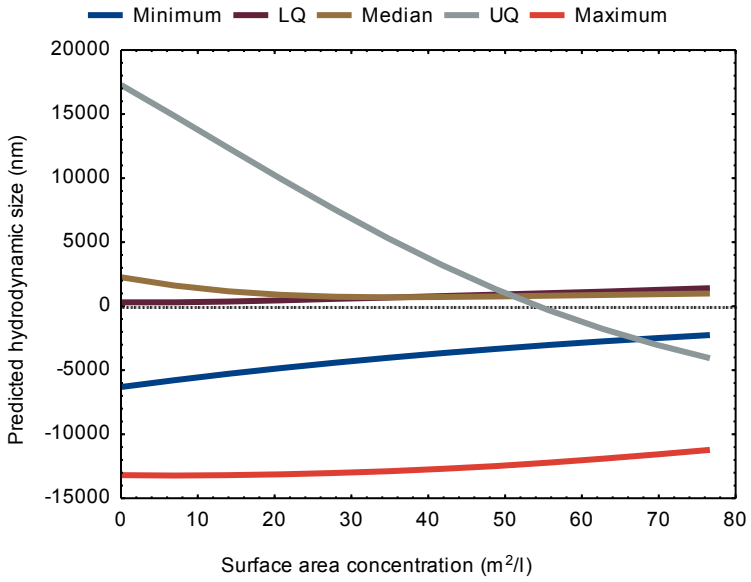


Figure 6.12 Contribution of the surface area concentration to predicting hydrodynamic size.

Independently, small primary size (Chowdhury et al. 2013) and the increase in concentration (Sillanpää et al. 2011) influence hydrodynamic size of $nTiO_2$ in the presence of organic matter. Integrated in the SAC, the primary size, and concentration of ENMs would not have similar influence on the demonstrated trends. Thus, the SAC requires further assessment to validate the estimated trends and support its application as an input in predictive modelling.

i) Initial Concentration of Organic Matter

The hydrodynamic size drops (LQ to UQ trends) with increasing initial concentration of organic matter (Fig. 6.13) agreeing with published literature (Keller et al. 2010; Ottofuelling et al. 2011; Chekli et al. 2015) because of stabilized electrophoretic mobility of ENMs (Domingos et al. 2009; Keller et al. 2010). The low and high concentration of organic matter induce fast and slow aggregation resulting in large but loose and small but compact sizes, respectively. The initial projected hydrodynamic size of the minimum and maximum trends conform to findings (Zhu et al. 2014). However, it is likely that other input variables contribute to the unexplained trends with increasing concentration of organic matter.

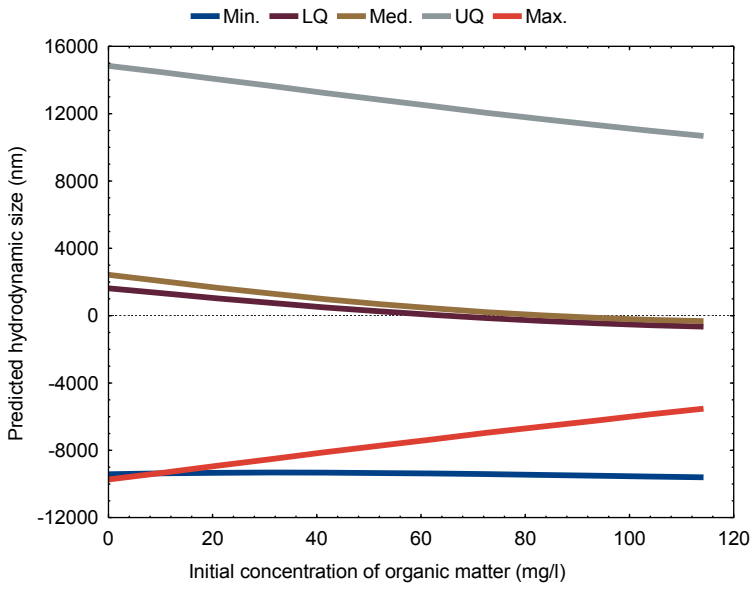


Figure 6.13 Contribution of the initial concentration of organic matter to predicting hydrodynamic size.

j) Composition of Anatase Crystals in $nTiO_2$

The hydrodynamic size increases in the positive and negative direction with an increase in the composition of anatase crystals (Fig. 6.14), suggesting a decrease in aggregation of predominantly rutile or high rutile / low anatase composition in $nTiO_2$ structure. The rutile phase has shown more stability in natural water than anatase (Iswarya et al. 2015) determined by different structural configurations (Zhou et al. 2013).

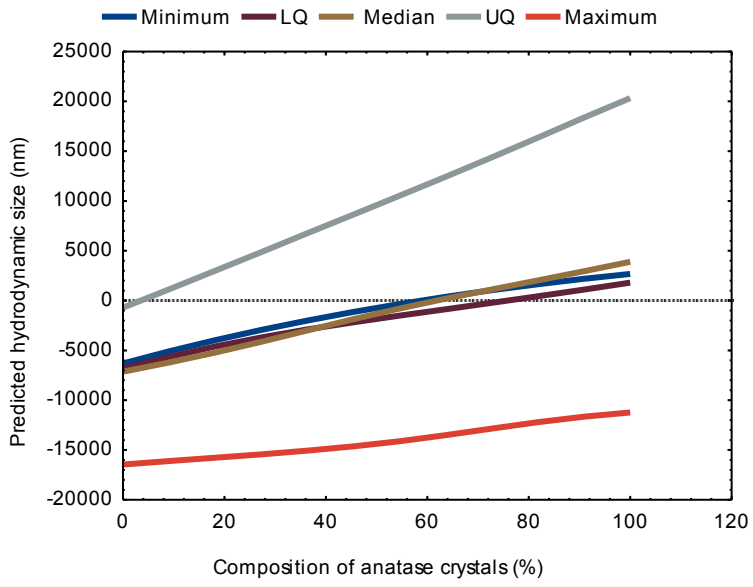


Figure 6.14 Contribution of percent anatase crystals to predicting hydrodynamic size.

Moreover, enhanced aggregation of mixed anatase / rutile nanoparticles (Liu et al. 2011; Iswarya et al. 2015) has been attributed to lower critical coagulation concentration than in predominantly anatase and rutile nanoparticles (Zhou et al. 2013). Thus, the role of mixed $nTiO_2$ crystals in additive effects on size evolution demonstrates potential application of crystal structure in predictive modelling of aggregation in aquatic environment.

The application of the ensemble model discussed in Chapter 5 to the data used to train the models presented in this Chapter, except for the surface coating, estimated organic adsorbates on $nTiO_2$. Figure 6.15 illustrates a nonlinear association between the predicted adsorbates and experimental hydrodynamic size.

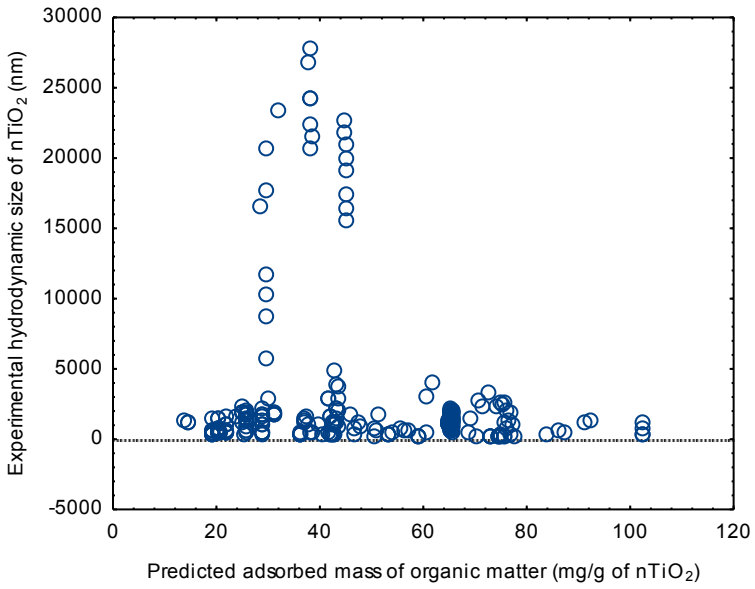


Figure 6.15 Association between predicted organic adsorbates on $nTiO_2$ and experimental hydrodynamic size of the ENM.

The lack of a specific pattern in the association suggests a broad dependence of size evolution of $nTiO_2$ on the quantity of adsorbed organic matter on the ENM. Hence, the surface coverage of adsorbates may stabilize and disperse ENMs or cause their aggregation and settlement in an aqueous environment.

6.4 Significance of the Model to Aquatic Environment

The model demonstrates the contribution of multiple physicochemical factors to aggregation and size evolution of $nTiO_2$, which in receiving aquatic systems would affect mobility, bioavailability, toxicity, and management issues of the ENM. The successful application of

both quantitative and qualitative data in the model shows possibilities of utilizing neural networks model in circumstances having multiple factors not well understood or characterized for general interpretations of the response.

The use of capping agents, organic / biological matter, and charged ions as qualitative data in predicting hydrodynamic size in static and semi-static aqueous systems has provided useful insights on similar phenomenon in analogous aquatic systems. Moreover, the model results show possibilities of applying the quantitative form of charged ions in predictive modelling. This approach can elucidate the influence of multiple ions' concentration and species on aggregation behaviour of ENMs that is challenging to measure empirically, and which is inferable to natural water systems. Besides, higher aggregate sizes from capped than uncapped $nTiO_2$ shows capability of the model to predict similar occurrence for capped ENMs used in nanoproducts and released into aquatic systems during their lifecycle.

The model permits making inferences about likely effects of dispersed and settled ENMs on water quality and ecosystem health because of aggregation. The qualitative property of highly dispersed ENMs is turbidity, which measures water quality (Lewis, 1996; Anderson, 2005; Bayram et al. 2012). Contingent on concentration-to-volume-ratios of $nTiO_2$ in an aqueous system, reported turbidity values of ≤ 2 to ≤ 150 nephelometric turbidity units (NTU) (Zhu et al. 2014) and 33-620 NTU (Botta et al. 2011) are higher than < 1 to ≤ 5 NTU allowable in drinking water (WHO 2011; SANS 2011).

The turbidity levels from incidences of ENMs discharged into aquatic systems enlightens the likelihood of constrained natural O_2 exchange between the water surface and aerobic organisms in the subsurface. Thus, the novelty of the MLP-NN model in predicting hydrodynamic size in an aqueous environment demonstrates its potential application as a decision-making tool in risk management of ENMs pollution, for instance, effecting water treatment to improve water quality.

6.5 Conclusion

Presented in this Chapter are results from an ensemble prediction of hydrodynamic size of capped and uncapped $nTiO_2$ using ten input variables drawn from properties of the ENM and aqueous environment. The findings show that all the input variables are relevant in determining hydrodynamic size in aqueous environment. Whereas capping agents, temperature, and ionic strength strongly influence size evolution of $nTiO_2$, the types of organic matter, duration of exposure, pH, and charged ions play transitional roles in defining

the stability of the ENM. Moreover, the initial concentration of organic matter, and crystal structure, primary size, morphology, and concentration of ENMs, display marginal but considerable contribution to aggregation behaviour.

The stability of capping agents in an aqueous environment is a good indicator of size evolution owing to likely assemblage of degraded agents with organic ligands. Moreover, application of properties of generic water systems as model inputs can over- and under-estimate likely hydrodynamic size of ENMs in an aquatic environment, respectively, depending on system variables. Simulated predictions of hydrodynamic size reveals a diverse influence of input variables that agrees as well as contrasts findings in published literature. The disparities could be model uncertainties brought about by the use of comparable response data generated under different exposure conditions.

A nonlinear relationship between adsorption and aggregation behaviour necessitates case-by-case assessment of adsorbates' effects on size evolution of ENMs in an aquatic environment where influential factors considerably vary. Accordingly, a comprehensive characterization of ENMs properties that affect inner- and outer sphere adsorption can support predictive modelling. The models developed may then explain potential associations between adsorption and aggregation and resultant effects that would be relevant in an aquatic environment.

The response data used in this study were an average of uni- or multi-modal distribution of hydrodynamic sizes. Thus, the successful prediction of the sizes may be challenging to project in an aquatic environment because of expected nanoparticle dispersity. However, the results demonstrate the capability of MLP-NN predictions inspiring development of mechanistic models that approximate appropriate hydrodynamic sizes of ENMs in an aquatic environment. These scenario can be supported by combining multiple known variables in the aquatic system to predict the response.

Chapter 7 - Prediction of Toxicity Effects of $nTiO_2$ on Algae and *Daphnia magna*

7.1 Summary

This Chapter presents research findings that address the third and final part of the second objective and the second and last part of the third objective. First, multilayer perceptron neural networks model is used to predict inhibition of nano-titanium dioxide ($nTiO_2$) on biomass growth in algae and reproduction in water flea, *Daphnia magna* as independent toxicity responses. Second, application of models implemented in Chapters 5 and 6 to fit input toxicity-based data separately approximated adsorbed mass of organic matter on $nTiO_2$ and hydrodynamic size of ENMs, respectively. The implemented models answer the second and last part, and the third question of the research questions presented in Section 1.6 of Chapter 1, which sought to establish relationships between causal factors and responses, and adsorption and aggregation behaviour of $nTiO_2$ and corresponding toxicity effects.

The model inputs include surface area concentration (SAC), composition of anatase crystals, and surface coating as properties of $nTiO_2$; and pH, ionic strength, temperature, duration of exposure, initial total carbon content and charged ions as abiotic factors. Additionally, types of organic matter (abiotic), and organism culture and species (biotic) are inputs specific for estimating inhibited biomass growth in algae. Besides, organism growth phase, feeding, and generation are specific biotic factors used to estimate inhibited reproduction in *D. magna*.

The ensemble models selected based on validation set errors predicted inhibited biomass growth in algae and reproduction in *D. magna* with mean square errors of 0.522 and 0.453, and $R^2 = 0.525$ and $R^2 = 0.536$, respectively. The predictions show sufficient approximation of experimental data. The algae species, and *D. magna's* generation and feeding are leading causal factors linked to aquatic food web and trophic transfer of ENMs. Whereas the surface chemistry of $nTiO_2$ influences toxicity on algae more than SAC, the reverse is applicable to toxicity on *D. magna*. The surface capping agents reduce toxicity effects on both organisms.

The exposure medium influences toxicity outcomes because of the interactions between $nTiO_2$ and diverse nutrients, which surmises similar effects in aquatic settings. The response of algae species to $nTiO_2$ is direct considering higher ranking than organism culture. Moreover, the presence of natural organic matter can mitigate toxicity effects on algae. In

both models, the temperature and pH depict very low sensitivity because of the narrow range of values in the datasets. However, high initial projected inhibitions suggest a constant variation rather than no influence on toxicity by the two variables.

The relationship between adsorption off organic matter on $nTiO_2$, aggregation, and size evolution of $nTiO_2$, and toxicity effects on algae and *D. magna* is nonlinear. Whereas a decrease in adsorption and increase in aggregation play a major role in toxicity on algae, a decrease in adsorption and aggregation have similar effects on *D. magna* toxicity. The findings illustrate the potential application of MLP-NN in nanoecotoxicology research that may in the longer-term save time and money in conducting specific experimental analyses.

7.2 Introduction

Advancement in the use of ENMs in nanoproducts and their potential release into the aquatic environment has necessitated testing their ecotoxicity effects. Dose-response acute and chronic toxicity tests conducted using organisms, which involve cells, generated training data applied in this study. Effective concentration $\left(EC_{(x)}\right)$ - based toxicity endpoints, where (x) is a percent effect, were measured using standard ecotoxicology testing protocols adopted in nanoecotoxicology (Appendix 4.4-4.5). Dose-response models fitted to toxicity data, for example, log-logistic, log-normal, Weibull-1 and Weibull-2, estimate the degree of chemical hazards for selected endpoints (Ritz 2010). These models generate different distribution functions of $EC_{(x)}$ whose limitations as a measure of toxicity have been articulated (Jager 2011). More importantly, the lack of synergy between dynamic environmental factors and EC_x limits inference of findings to explain related phenomena in aquatic ecosystems (Jager 2011).

Environmental abiotic and biotic factors promote survival of organisms in ecosystems and govern the behaviour kinetics of ENMs reported in Chapters 5 and 6, and in turn toxicity effects discussed in this Chapter. In the dynamic aquatic environment, changes in any of the influential factors determine the extent of toxicity effects. The structure and heterogeneity of laboratory derived data used simulate complex interactions between ENMs and environmental factors influencing toxicity effects on algae and *D. magna* as indicator organisms. Input variables applied to train models reported in Chapters 5 and 6, and five designated biotic factors, are used to estimate toxicity responses discussed in this Chapter. In principle, adsorption and aggregation kinetics influence toxicity effects of ENMs on organisms.

Here, properties of ENMs and environmental factors are summarily introduced in connection with organisms. While damage to algae by anatase impairs cell membrane and DNA, and rutile inhibits photosynthesis, their combined effects induce structural impairment that hinders CO_2 uptake and subsequent metabolism (Iswarya et al. 2015). Whereas SAC avails chemical reaction sites of ENMs to organisms, ionic strength controls ions' activity and ENMs-organism charges at reaction sites. Besides, the temperature and pH determine the uptake of toxicants, CO_2, O_2, and consequent metabolic reactions. A recent study (Kim et al. 2016) elucidates the influence of duration of exposure on growth and development of organisms, and partitioning, accumulation and trophic transfer of ENMs in ecosystems. Total carbon provides nutrition to organisms and acts as a barrier to their interaction with ENMs, thus, control the degree of toxicity effects. Section 4.4 and 4.5 in Chapter 4 describes details about these input variables.

Nominal levels of surface coating variables used in this study include uncoated (bare) ENMs as well as those coated with Al_2O_3 (Al); Al_2O_3-SiO_2 (AlSi); $Al(OH)_3$ and polydimethylsiloxane (AlPDMS); and $Al(OH)_3$, PDMS, glycerin and other organic compounds (AlPDMSGI). The data used in this study were generated using assorted standard synthetic water and mesocosm freshwater (MesoF) (Keller et al. 2010). The chemistry of algal test media (Algal TM) (OECD 201, 2011), MesoF, Elendt 4 (Elendt and Bias 1990), reconstituted *D. magna* culture media (RDCM) (OECD 211, 2012), hard water (HW) and moderately hard water (MHW) (USEPA 2002) represent mixed charged ions evaluated in this study. Moreover, natural organic matter (NOM) and total carbon (TC) from spiked electrolytes define levels of the types of organic matter limited to biomass growth response.

Molecular mechanisms of ENMs ecotoxicity are not well understood notwithstanding photocatalytic oxidative stress linked to the generation of reactive oxygen species (ROS) as a defense mechanism (Reeves et al. 2008; Kim et al. 2010; Metzler et al. 2011). Still, a low solubility of $nTiO_2$ implies possible physical rather than ionic uptake pathway. Once in contact with chemical pollutants, organisms have different behavioural, physiological and biochemical mechanisms of coping with the related stress, which leads to immobilization or inhibition as defense mechanisms. It has been suggested that algae overgrowth is a defense mechanism against toxicants (Aguilera et al. 2002). The inhibition of biomass growth in algae and reproduction in *D. magna* is a physiological response after exposure to ENMs.

Derived from measured toxicity endpoints (refer to Section 4.7.1 in Chapter 4), the percent toxicity effect adopted in this study deviates from the EC_x - based indicator to provide a basis

for measuring the increase and decrease in population levels during exposure. The approximated inhibited biomass growth involves twelve algae species (Table 4.8), exposed in axenic and mixed cultures (Section 4.6.1 and 4.7.2) as categorical values of two biotic factors. Moreover, Section 4.6.1 and Table 4.9 outline three growth phases, six organism generations, and six types of food investigated as nominal levels of biotic factors instrumental to ENMs' toxicity effects on *D. magna*. The growth phase is an index that estimates population levels of surviving organisms described by the discrete time Euler-Lotka equation (Eqn. 7.1).

$$\int_{\alpha}^{\beta} e^{-rx}\, \iota(x).m(x).dx = 1 \tag{7.1}$$

where x is the age of individuals, r is the population growth rate, $\iota(x)$ is the specific survival probability to age x, $m(x)$ is the female fecundity at age x, β is the last reproduction age and α is the first reproduction age. For a given toxicant exposure period, individual growth, survival, and fecundity of organisms are intermediate endpoints to population levels. Because of inadequate information or contradictory reports, potential input factors excepted in this study include growth phase and carbon reserves in algae, pH variations owing to ENMs and metabolic reactions, and effects of irradiation.

The association between behaviour kinetics and the derived toxicity responses is estimated and reported towards the end of the Chapter. To the best of the author's knowledge, this is the first time neural networks is applied to estimate the toxicity of ENMs to aquatic organisms using integrated data and relative toxicity as a new response variable. The relative increase and decrease in toxicity effects concern potential population dynamics of algae and *D. magna* owing to ENMs toxicants and the impact on energy transfer in the ecosystem.

7.3 Results and Discussion

The modelling method and procedures outlined in Chapters 5 and 6 apply in this Chapter to develop predictive ensemble models, assess sensitivities of inputs, and simulate input-output associations. Two data subsets used comprise 12 input variables, and 226 and 154 training examples for inhibited biomass growth in algae and reproduction in *D. magna*, respectively. A split of the data into 90% for parallel training (80%) and verifying the accuracy (20%) of models withheld 10% for out-of-sample testing of the models. The initial seeds for randomly splitting the data during parallel training were 1, 56, 1,234, 4,986, and

9,700; and 14, 135, 2,780, 4,694, and 7,980 for inhibited biomass growth in algae and reproduction in *D. magna*, respectively. The mean square error (MSE) was a measure of the models' performance based on validation set (implementation) and testing set (prediction).

The tanh and identity were set as input and output activation functions. The maximum number of iterations was set at 1,000. Network configurations were 8, set between 4 and 11, inclusive, and 5 models retained after every training experiment. Initial model training using fixed 1, 234, and 2,780 subsampling seeds, and 7 hidden neurons established weight decay (α) values applied in subsequent model development (Table 7.1). The smallest validation set error determined the α value implemented (highlighted in grey).

Table 7.1 Weight regularization values for toxicity-based models

Weight decay (α)		Standalone MSE		Ensemble MSE
Hidden layer	Output layer	Validation	Test	Test
Inhibited biomass growth in algae				
10^{-7}	10^{-8}	1.476 ± 3.512	1.830 ± 1.048	0.780
10^{-6}	10^{-7}	0.266 ± 0.146	0.720 ± 0.201	0.570
10^{-5}	10^{-6}	0.226 ± 0.090	0.741 ± 0.090	0.670
10^{-4}	10^{-5}	0.245 ± 0.076	0.637 ± 0.109	0.610
10^{-3}	10^{-4}	0.245 ± 0.079	0.597 ± 0.095	0.591
10^{-2}	10^{-3}	0.240 ± 0.075	0.571 ± 0.067	0.550
10^{-1}	10^{-2}	0.324 ± 0.086	0.741 ± 0.090	0.731
Inhibited reproduction in *D. magna*				
1×10^{-2}	1×10^{-3}	0.210 ± 0.050	0.490 ± 0.040	0.474
8×10^{-3}	8×10^{-4}	0.223 ± 0.023	0.527 ± 0.142	0.491
6×10^{-3}	6×10^{-4}	0.226 ± 0.045	0.577 ± 0.115	0.528
4×10^{-3}	4×10^{-4}	0.240 ± 0.060	0.587 ± 0.132	0.545
2×10^{-3}	2×10^{-4}	0.237 ± 0.053	0.594 ± 0.170	0.551
1×10^{-3}	1×10^{-4}	0.275 ± 0.390	1.124 ± 1.000	0.840

7.3.1 Results for Inhibited Biomass Growth in Algae

Twenty-five models retained for each of the 8 network topologies were assessed to reveal a narrow influence of the number of hidden neurons on performance presented in Table 7.2.

The average performance of standalone models developed was comparable and marginally improved by 1.7-3.4% after aggregation (Table 7.2). Figures 7.1 and 7.2 present the validation and out-of-sample set error variances of ensemble models, respectively. The ensemble of 25 standalone models trained using 7 hidden neurons was implemented because of exhibiting the minimum validation set error (Table 7.2) (in grey).

Table 7.2 Summary of standalone and ensemble model performance

	Standalone model MSE (0.95 confidence interval)			Ensemble model MSE		Ensemble
Topology	Training	Validation	Testing	Testing	Reduction (%)	R^2
MLP 10-4-1	0.255 ± 0.009	0.250 ± 0.021	0.533 ± 0.020	0.515	3.4	0.543
MLP 10-5-1	0.252 ± 0.006	0.242 ± 0.020	0.528 ± 0.021	0.515	2.4	0.538
MLP 10-6-1	0.246 ± 0.006	0.241 ± 0.020	0.534 ± 0.023	0.522	2.2	0.524
MLP 10-7-1	0.245 ± 0.006	0.237 ± 0.020	0.531 ± 0.020	0.522	1.7	0.525
MLP 10-8-1	0.246 ± 0.008	0.247 ± 0.024	0.536 ± 0.026	0.520	3.0	0.526
MLP 10-9-1	0.243 ± 0.005	0.240 ± 0.020	0.532 ± 0.025	0.522	1.9	0.525
MLP 10-10-1	0.244 ± 0.007	0.237 ± 0.021	0.525 ± 0.023	0.514	1.9	0.530
MLP 10-11-1	0.241 ± 0.006	0.237 ± 0.021	0.536 ± 0.023	0.528	1.7	0.516

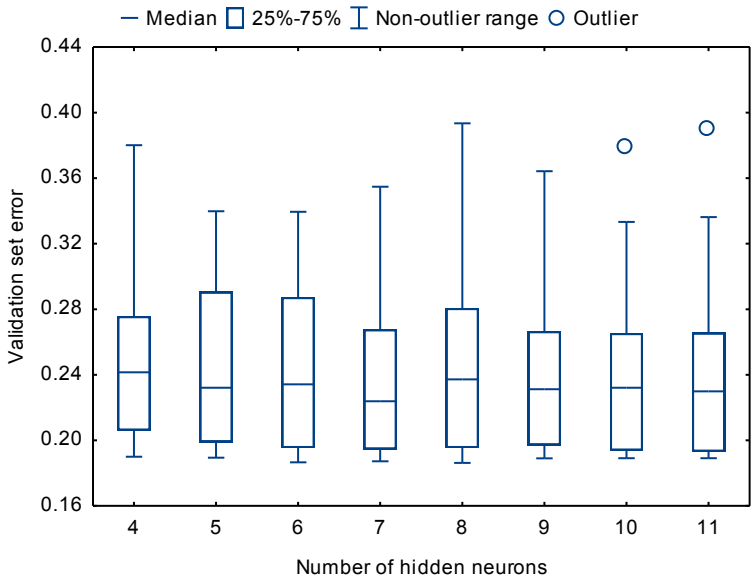


Figure 7.1 Average validation set errors of networks having a different number of hidden neurons.

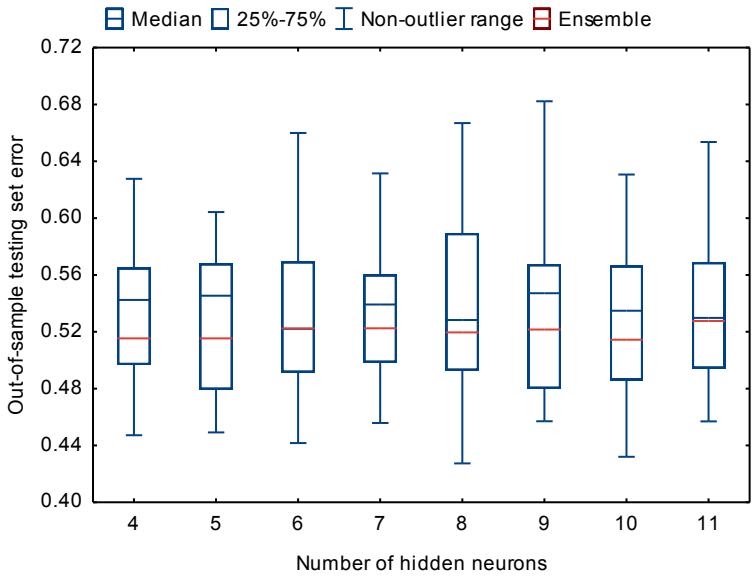


Figure 7.2 Testing set errors of networks having a different number of hidden neurons.

The MSE of the selected ensemble model based on 90% training and validation data sets was 0.480. A scatter comparison of experimental and predicted values shows sufficient estimation of inhibited biomass growth in algae with $R^2 = 0.525$ (0.95 prediction interval) (Table 7.2 and Fig. 7.3).

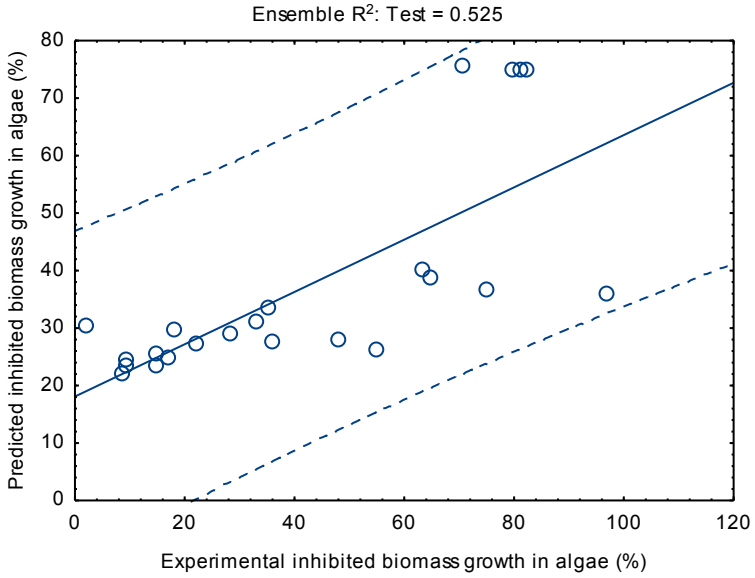


Figure 7.3 Ensemble prediction of inhibited biomass growth in algae.

A predicted versus residual inhibited biomass growth scatter comparison shows the independence ($R^2 = 0.02$ and slope of -0.87) and under- / over-prediction of the response

without a noticeable pattern (0.95 prediction interval) (Fig. 7.4) (*Left*). The normality ($W = 0.930$) of residual values is unbalanced and displays a right-skewed distribution (Fig. 7.4) (*Right)* with a mean, median and standard deviation of 4.88, 2.11 and 20.42, respectively.

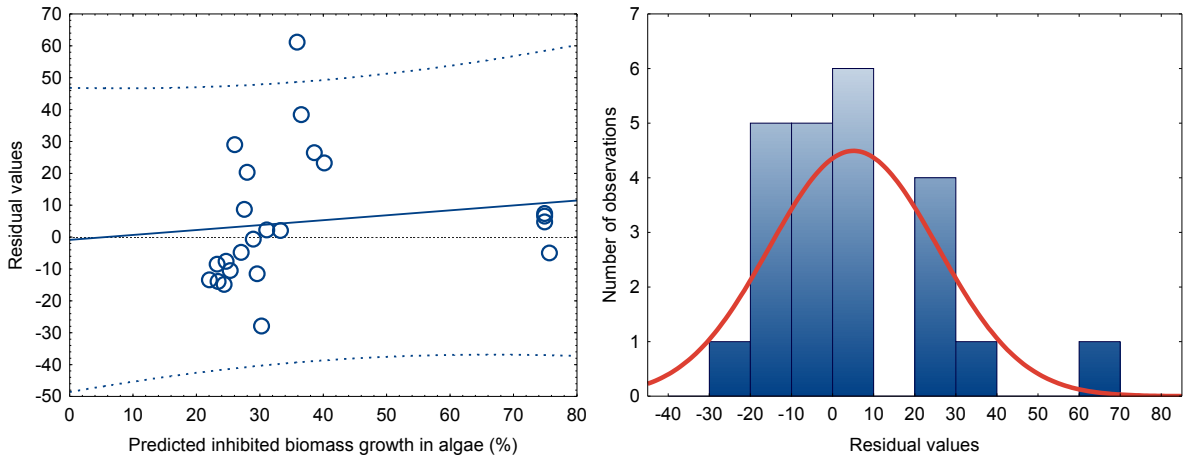


Figure 7.4 Predicted inhibited biomass growth versus residual values (*Left*) and normality of residual values (*Right*).

7.3.3.1 Relative Importance of Input Variables

The previous Chapters outline procedures for evaluating sensitivities of input variables and their contribution to the response that are repeated in this Chapter. The MSE computed by perturbing input variables in the fitted ensemble model ranges from the smallest (0.478) to the largest (0.720) for pH and surface coating, respectively (Fig. 7.5) (in italics). The computed sensitivities represent relative importance of input variables in predicting inhibited biomass growth in algae (Fig. 7.5) (in bold).

All input variables exert various proportions of influence on biomass growth except for the pH and temperature, likely because of their narrow range of values that are insensitive to the model. The most influential variables are; surface coating, organism species, charged ions, organism culture, and types of organic matter in the given order (Fig. 7.5). The surface chemistry ranks higher than the size, concentration, morphology and crystal structure of ENMs in elucidating biomass growth inhibition. The importance of organism species exceeds that of organism culture by 1.32-fold, interpreted as greater direct than associative effects of $nTiO_2$ to algae during exposure. Table 7.3 presents independently computed MSE and percent sensitivities (in italics) of each nominal level of the categorical variables.

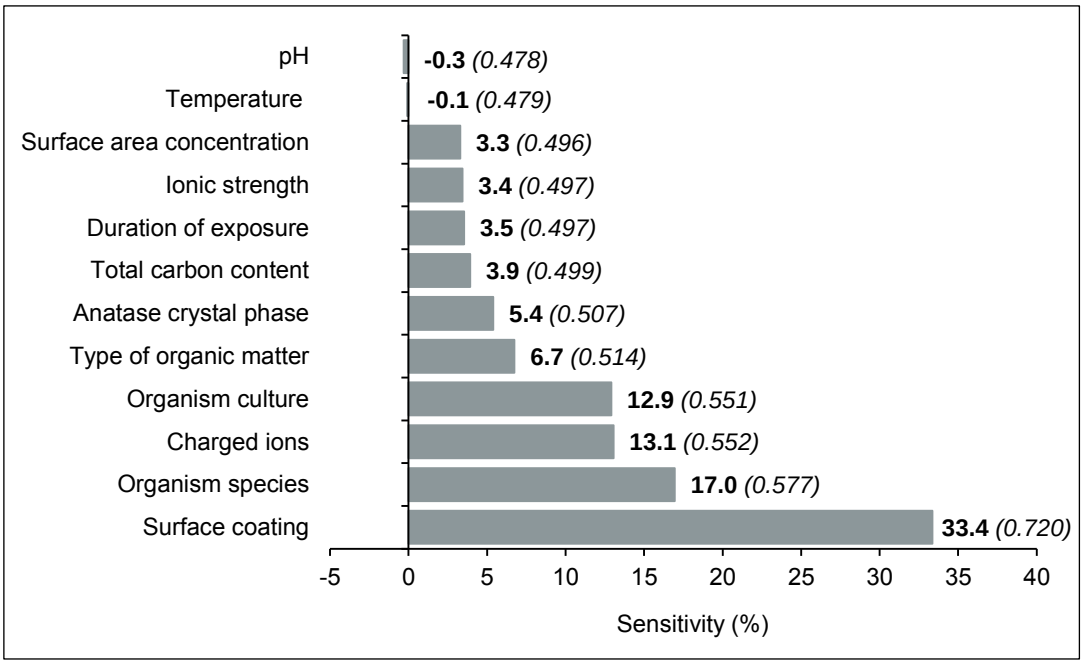


Figure 7.5 Sensitivity of input variables in predicting inhibited biomass growth in algae.

Table 7.3 Sensitivity of nominal levels in predicting inhibited biomass growth in algae

Input variable Algae species	MSE and sensitivity	Rank	Input variable Surface coating	MSE and sensitivity	Rank
Tabularia fasciculata	0.638 (*24.8*)	1	Alumina and silica (AlSi)	0.830 (*42.2*)	1
Anabaena spp.	0.630 (*23.8*)	2	Alumina (Al)	0.826 (*42.0*)	2
Stigeoclonium tenue	0.625 (*23.3*)	3	Bare	0.503 (*4.6*)	3
Planothidium lanceolatum	0.610 (*21.3*)	4	Charged ions		
Spyrogyra communis	0.592 (*19.0*)	5	algal test media (Algal TM)	0.570 (*15.7*)	1
Navicula subminiscula	0.578 (*17.0*)	6	Mesocosm freshwater (MesoF)	0.534 (*10.2*)	2
Pseudokirchneriella subcapitata	0.568 (*15.6*)	7	Organism culture		
Oscillatoria spp.	0.558 (*14.1*)	8	Axenic	0.568 (*15.6*)	1
Selenastrum minutum	0.544 (*12.0*)	9	Mixed	0.533 (*10.0*)	2
Chlorella vulgaris	0.533 (*10.0*)	10	Types of organic matter		
Scenedesmus quadricauda	0.528 (*9.2*)	11	Total carbon (TC)	0.523 (*8.3*)	1
Chlamydomonas moewusii	0.525 (*8.7*)	12	Natural organic matter (NOM)	0.505 (*5.1*)	2

The data presented in Table 7.4 simulates theoretical exposure conditions whose minimum, lower quartile (LQ), median, upper quartile (UQ) and maximum values in the dataset were used to evaluate model sensitivity to continuous changes in input variables. The spiking of NOM in standard testing media in nanoecotoxicology studies (Cupi et al. 2015) validates the simulated minimum to median exposure conditions. The values of input variables are repeated except for the SAC and duration of exposure.

Table 7.4 Hypothetical exposure conditions to project inhibited biomass growth in algae

Input variable	Minimum	LQ	Median	UQ	Maximum
Surface area concentration (SAC) ($m^2/ℓ$)	5.8E-4	2.50	5.00	15.00	80.64
Temperature (°C)	18.00	18.00	18.00	22.00	23.80
Initial carbon content (mg/ℓ)	2.15	7.16	63.45	63.45	63.45
Duration of exposure (hours)	24.0	72.0	144.0	288.0	600.0
Types of organic matter	NOM	NOM	NOM	TC	TC
Charged ions	Algal TM	Algal TM	MesoF	MesoF	MesoF
Surface coating	Bare	Bare	Bare	Bare	AlSi
Composition of anatase crystals (%)	0.0	82.0	82.0	82.0	100.0
pH	7.45	7.75	8.38	8.38	8.38
Ionic strength (mM)	1.00	1.67	7.18	7.18	7.18
Organism culture	Axenic	Axenic	Mixed	Mixed	Mixed
Organism species	*P. subcapitata*	*P. subcapitata*	*C. moewusii*	*Oscillatoria spp.*	*T. fasciculata*

With reference to Figure 7.5, and Tables 7.3 and 7.4, Sections (a) to (j) focus discussions on the contribution of input variables to predicting inhibited biomass growth in algae and likely implications. Notably, linear profiles of continuous variables versus the response is a characteristic of screening out caused by the regularization parameter. The AlSi capping on $nTiO_2$ likely explains the consistent low inhibitions estimated using maximum data. Remarks by Das et al. (2013) link reduced bioavailability and toxicity effects of ENMs to capping agents. Moreover, the LQ data estimated higher toxicity of $nTiO_2$ on *P. subcapitata* than minimum data, perhaps because of the following factors discussed in respective Sections,

- A 3.0-fold, 3.3-fold, and 4,310-fold increase in duration of exposure, carbon content, and SAC (due to increase in concentration of smaller particles of $nTiO_2$), respectively.

- A 4.0% and 67.0% increase in the pH and ionic strength, respectively, and presence of anatase crystals.

The *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C. moewusii*$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$, and *T. fasciculata*$_{(max}$ are designates that differentiate toxicity of $nTiO_2$ on the algae species estimated using the minimum, LQ, median, UQ and maximum data values, respectively.

a) Surface Coating

The individual sensitivity of AlSi is 1.01-fold and 9.24-fold higher than that of Alumina and bare $nTiO_2$, respectively (Table 7.3). However, the sequence AlSi < Alumina < bare (Fig. 7.6) inverts the sensitivities in predicting the response. The capping agents decrease toxicity of $nTiO_2$ on *P. subcapitata*$_{(LQ)}$, *P. subcapitata*$_{(m)}$, *T. fasciculata*$_{(max)}$, *Oscillatoria spp.*$_{(UQ)}$, and *C. moewusii*$_{(med.)}$ by 1.55-fold, 1.94-fold, 1.90-fold, 1.96-fold, and 3.17-fold, respectively.

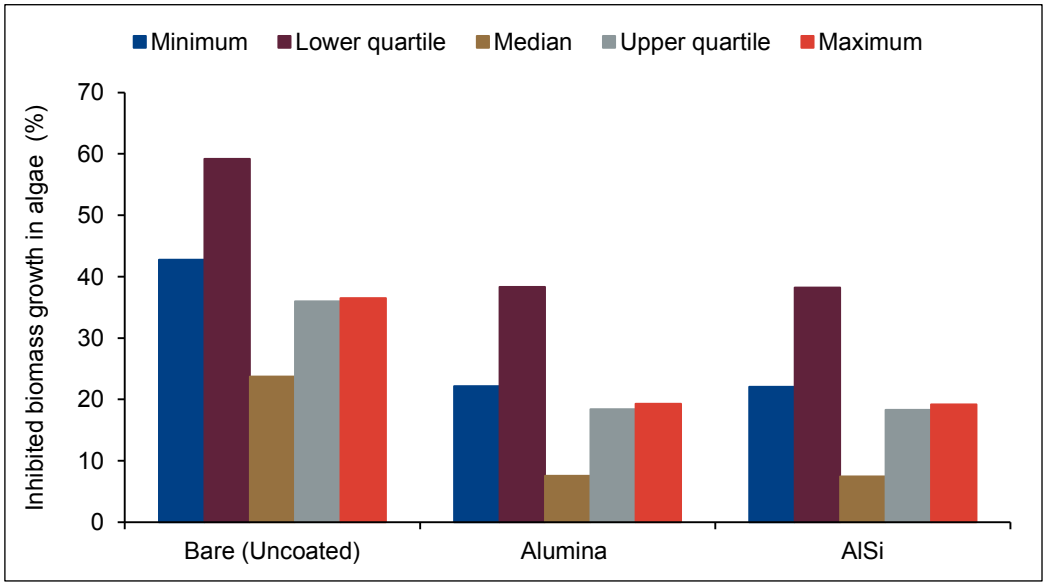


Figure 7.6 Influence of ENMs' surface coating to inhibiting biomass growth in algae.

The particle-particle and particle-solid interactions of bare ENMs would promote clustering, thus, bioavailability as dispersed on settled solids. In contrast, potential hydrolyzation of SiO_2 and amphoteric reactions of Al_2O_3 may alter the surface chemistry and enhance the reactivity of ENMs leading to toxicity effects. Similar processes are likely to occur in aquatic systems, thus, the appropriateness of using this variable to estimate toxicity effects.

b) Algae Species

The individual sensitivity is highest for *T. fasciculata* and lowest for *C. moewusii* (Table 7.3).

Moreover, a 2.3% - 16.0% difference between preceding sensitivities of intermediate algae species suggests diverse physiological reactions to ENMs that affects biomass growth.

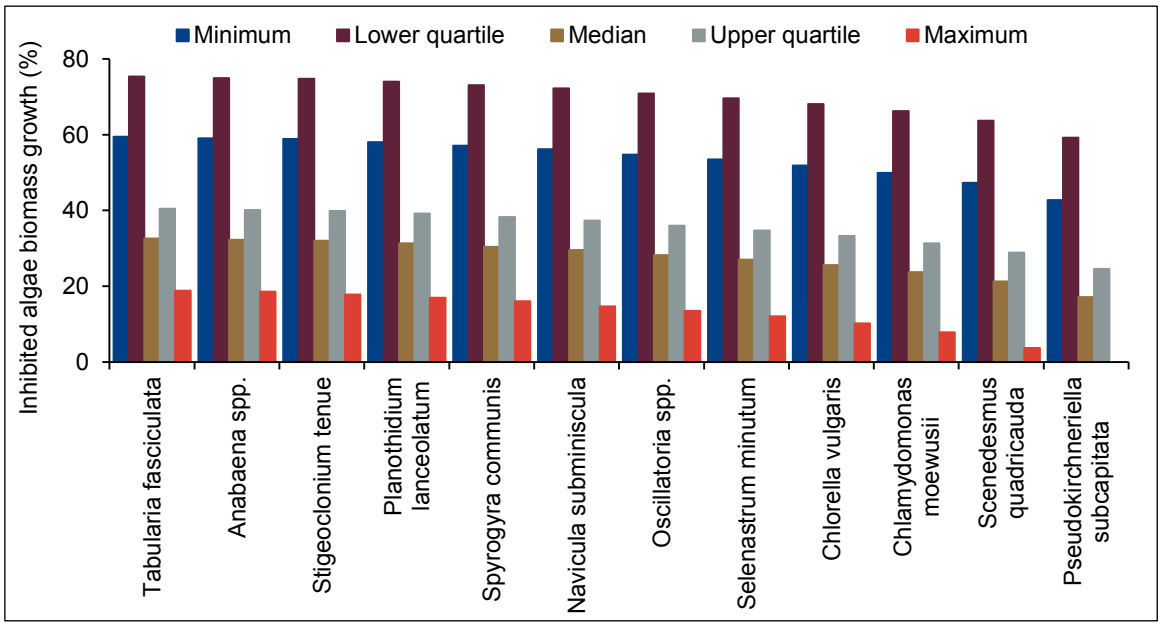


Figure 7.7 Influence of algae species on inhibited biomass growth in algae.

The estimated inhibitions decrease in the LQ > minimum > UQ > median > maximum sequence indicating toxicity effects determined by habitats and algae species. Algae species thrive in diverse aquatic ecosystems consisting a multitude of nutrients and other growth-promoting or limiting factors. Under all simulated exposure conditions, the sequence of toxic response to algae is 1 > 2 > 3 > 4 > 5 > 6 > 8 > 9 > 10 > 12 > 11 > 7 (refer to the ranks listed in Table 7.3 for identity of algae species) (Fig. 7.7). The diverse metabolisms of algae species determine their toxicity reactions during exposure (Cardinale et al. 2012). It appears impractical to extrapolate toxicity findings using model algae species, thus, the need to utilize modelling as a tool to support nanoecotoxicology risk evaluation.

c) Charged Ions and Organism Culture

The sensitivity of charged ions (Fig. 7.8) is higher in monocultures (15.6%) exposed in Algal TM (15.7%) than in mixed cultures (10.0%) exposed in MesoF (10.2%) (Table 7.3). The speciation of charged ions in MesoF influences the stability and bioavailability of ENMs that acquire high negative surface charges (Keller et al. 2010), hence, reduced toxicity effects. The $nTiO_2$ acquire a Under all exposure conditions, projected inhibitions decrease in the sequence *P. subcapitata*$_{(LQ)}$ > *Oscillatoria spp.*$_{(UQ)}$ > *P. subcapitata*$_{(m)}$ > *C. moewusii*$_{(med.)}$ > *T. fasciculata*$_{(max)}$ (Fig. 7.8).

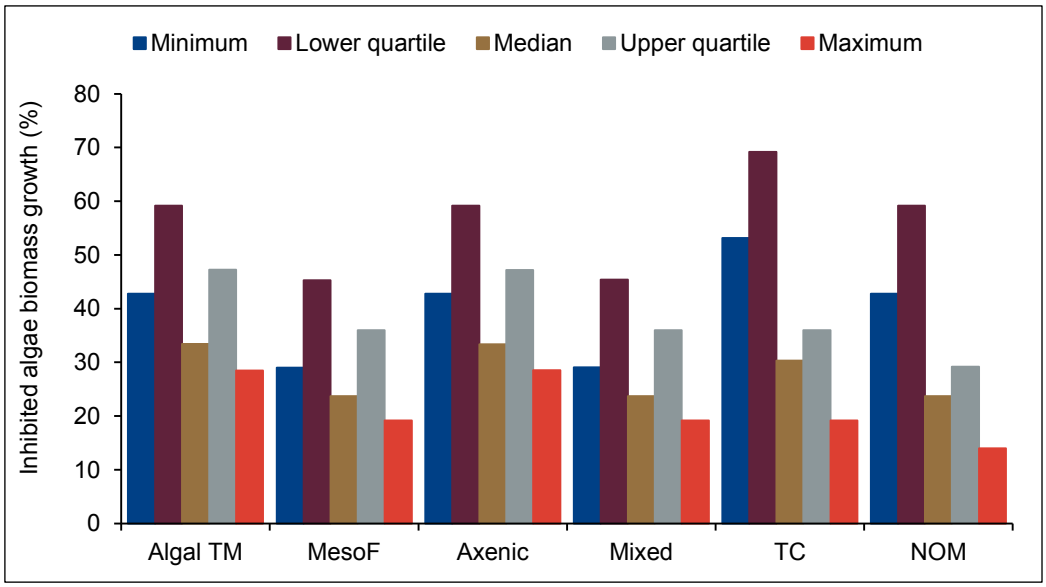


Figure 7.8 Influence of charged ions and organism culture on inhibited biomass growth in algae.

The correlation between charged ions and organism culture imply likely overestimation of toxicity to single algae species exposed in standard media, which may be irrelevant to the aquatic systems. Thus, projections in MesoF suggest moderated toxicity effects of $nTiO_2$ in aquatic systems comprising heterogeneous organism communities and ecological variables.

d) Types of Organic Matter

The sensitivities of TC (8.3%) and NOM (5.1%) (Table 7.3) and Figure 7.8 are similar, indicating a higher toxicity in the presence of the former than latter. The magnitude of toxicity effects decreases in the sequence: *P. subcapitata*$_{(LQ)}$ > *P. subcapitata*$_{(m)}$ > *Oscillatoria spp.*$_{(UQ)}$ > *C. moewusii*$_{(med.)}$ > *T. fasciculata*$_{(max)}$ (Fig. 7.8). From a nourishment perspective, properties and speciation of carbon in spiked inorganic salts and in NOM found in pseudo-natural freshwater avail a direct and indirect nutrient reserve for algae metabolism, respectively. From a pollution standpoint, the complex nature of NOM and bridging adsorption with ENMs and various ions may reduce bioavailability and mitigate toxicity effects on algae compared to TC in standard media as observed in Figure 7.8.

e) Composition of Anatase Crystals

The biomass growth inhibition increases linearly with increasing composition of anatase $nTiO_2$ crystals (Fig. 7.9). The inhibitions on *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C. moewusii*$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$, *and T. fasciculata*$_{(max)}$ biomass growth increases by 54.3%, 58.1%, 201.0%, 99.2%, and 1,518.3% at 99.0% with increasing anatase,

respectively (Fig. 7.9). The trend agrees with findings (Ji et al. 2011; Clément et al. 2013; Campos et al. 2013; Mansfield et al. 2015). The independent and mixed effects of rutile and anatase crystals on algae (Iswarya et al. 2015) may cause additive and subtractive effects that elucidate the trends. Thus, this variable is worth exploring in predictive modelling.

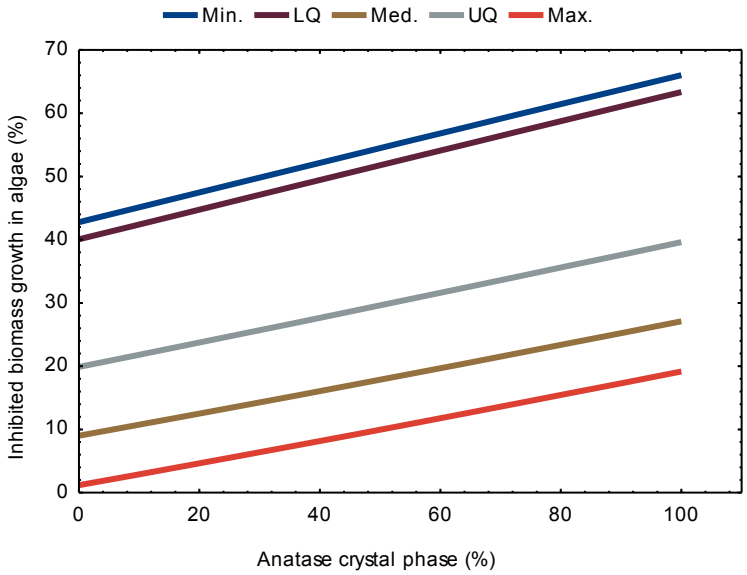


Figure 7.9 Influence of composition of anatase $nTiO_2$ crystals on inhibited biomass growth in algae.

f) Initial Total Carbon Content and Duration of Exposure

The toxicity of $nTiO_2$ on *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C. moewusii*$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$, *and T. fasciculata*$_{(max)}$ biomass growth decreases by 28.2%, 20.2%, 26.1%, 21.3%, and 29.6%, respectively, as the initial carbon content increases (Fig. 7.10). From the previous Chapters, potential promotion or obstruction of adsorption (Fig. 5.9) and reduction or enhancement of aggregation (Fig. 6.15) with increasing concentration of organic matter can, in turn, affect stability, bioavailability, and toxicity of $nTiO_2$ on algae species. Thus, a combination of adsorption and aggregation behaviour of $nTiO_2$ in the presence of carbon may avail dispersed or settled nanoparticles to algae, leading to diverse toxicity effects on organisms under different exposure conditions.

The projected trends show increasing inhibitions on *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C. moewusii*$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$, *and T. fasciculata*$_{(max)}$ biomass growth by 28.6%, 20.5%, 52.2%, 37.6%, and 149.5% with lengthy duration of exposure (Fig. 7.11), which agrees with findings (Aruoja et al. 2009;Ji et al. 2011). Comparatively, the previous Chapters showed

decreasing adsorption (Fig. 5.5) and aggregation (Fig. 6.12) with increasing duration of exposure, which would favour bioavailability of dispersed $nTiO_2$, thus, enhanced toxicity.

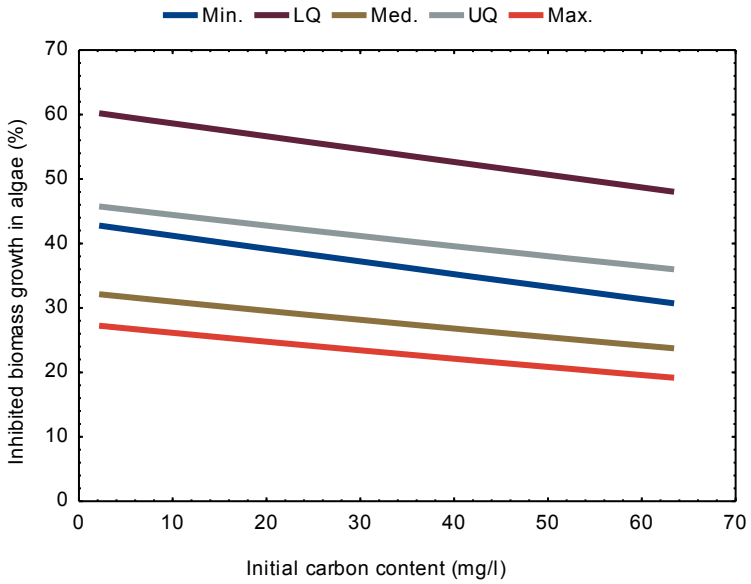


Figure 7.10 Influence of the initial carbon content on inhibited biomass growth in algae.

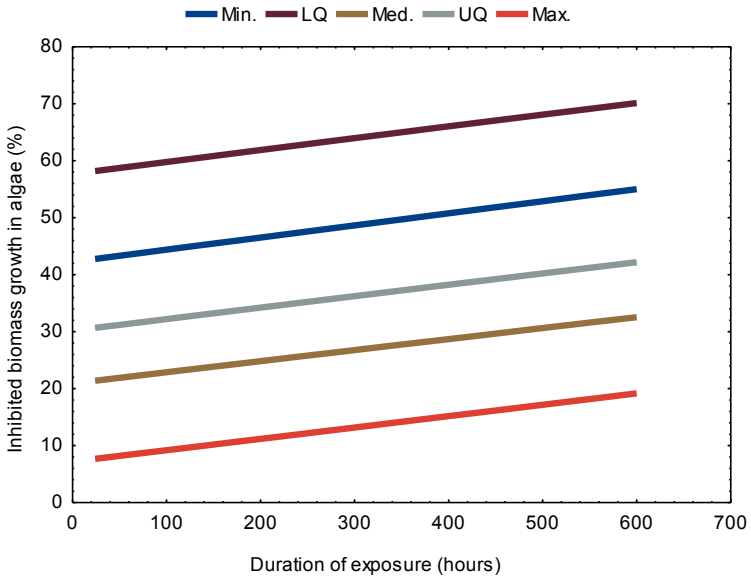


Figure 7.11 Influence of the duration of exposure on inhibited biomass growth in algae.

g) Ionic Strength

The initial inhibitions on *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C. moewusii*$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$, *and T. fasciculata*$_{(max)}$ biomass growth drop by 27.4%, 19.6%, 25.5%, 20.8%, and

29.0%, respectively, with increasing ionic strength (Fig. 7.12). The previous projections show that an increase in ionic strength increases adsorption (Fig. 5.12), increases or decreases aggregation (Fig. 6.7) that relate to the projections in Figure 7.12.

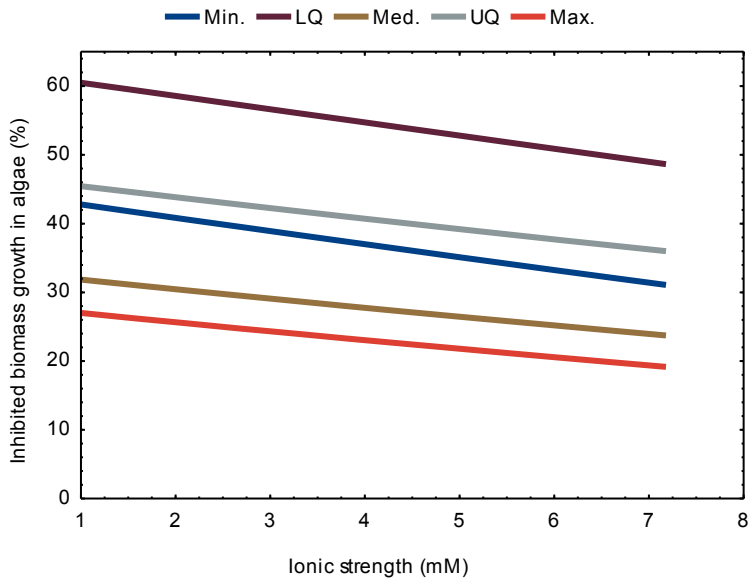


Figure 7.12 Influence of ionic strength on inhibited biomass growth in algae.

The trends demonstrate interface reactions not solely related to adsorption of organic matter since multiple ions coexist in the aqueous media used to generate data. At constant room temperature, activity coefficients of inorganic carbon species were found to decrease with increasing ionic strength (Goldman et al. 1971). Applying Eqn 4.11 in Chapter 4, constant temperature decreases activity coefficient of ions and increases the magnitude of zeta potential with increasing ionic strength. Thus, high values of zeta potential may increase electrostatic stabilization of ENMs, thus, surface reactivity, and reduce toxicity effects.

h) Surface Area Concentration

The SAC is a primary input variable in this study that requires an enabling environment for the effects to manifest. However, the low sensitivity of SAC (3.3%) in the model may be attributable to data uncertainties because of similar magnitudes of endpoints generated using diverse testing methods. The inhibitions on *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C. moewusii*$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$, *and T. fasciculata*$_{(max)}$ biomass growth decrease by 51.8%, 37.6%, 63.1%, 44.8%, and 47.6%, respectively, when SAC increases (Fig. 7.13).

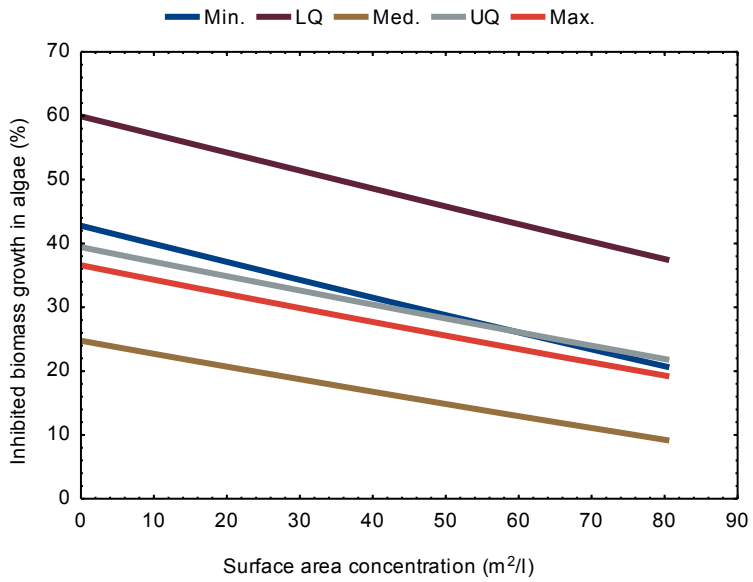


Figure 7.13 Influence of SAC on inhibited biomass growth in algae.

At very low NOM levels in purified water, toxicity of $nTiO_2$ to algae was found to increase with increasing concentration (Aruoja et al. 2009; Ji et al. 2011) and decreasing particle size (Hund-Rinke and Simon 2006). The findings from published literature do not explain the response trends extracted by the model with respect to SAC. Thus, it is essential to evaluate SAC as a predictor variable for potential application in nanoecotoxicology risk assessment.

i) Temperature and pH

The -0.1% sensitivity of temperature (Fig. 7.5) seems related to *ca.* constant variations rather than negligible effects on toxicity of $nTiO_2$ on algae (Fig. 7.14). The initial inhibitions (20.0% to 54.0%, no order) on *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C. moewusii*$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$, *and T. fasciculata*$_{(max)}$ biomass growth increased by 6.0%, 4.1%, 3.2%, 3.7%, and 2.0%, respectively, with increasing temperature (Fig. 7.14). The aquatic temperature (refer to Table 4.7 in Chapter 4) is spatiotemporal in view of the metabolic and thermodynamic processes, and seasonal variations, that would affect the behaviour of ENMs discussed in previous Chapters. However, a 5.8°C range of values (Table 7.4) may explain negligible model sensitivity to temperature despite showing high initial toxicity outputs.

The pH is the least sensitive variable (-0.3%) (Fig. 7.5) that shows a trivial rise and fall in the projected inhibitions of $nTiO_2$ on algae biomass growth (Fig. 7.15). An increase in pH decreases toxic biomass growth inhibitions of *P. subcapitata*$_{(m)}$, *P. subcapitata*$_{(LQ)}$, *C.*

moewusii$_{(med.)}$, *Oscillatoria spp.*$_{(UQ)}$ by 5.1%, 3.8%, 0.3%, and 2.1%, respectively, and increases that of *T. fasciculata*$_{(max)}$ by 1.3% (Fig. 7.15).

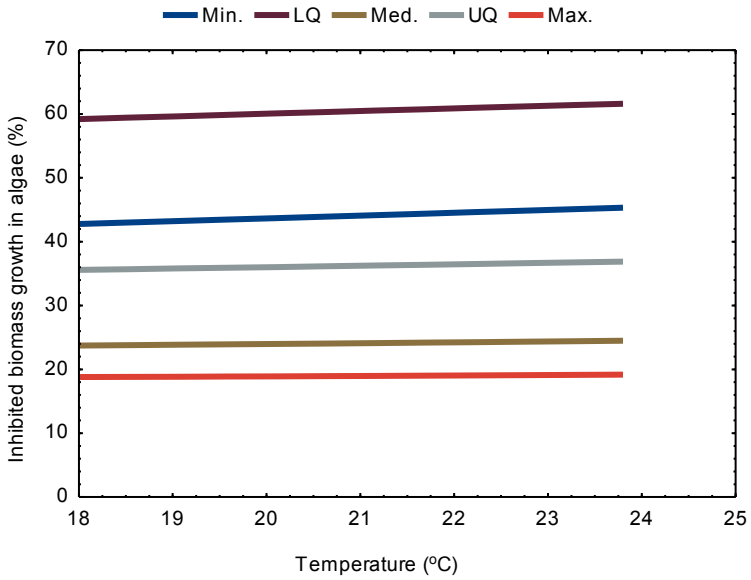


Figure 7.14 Influence of the temperature on inhibited biomass growth in algae.

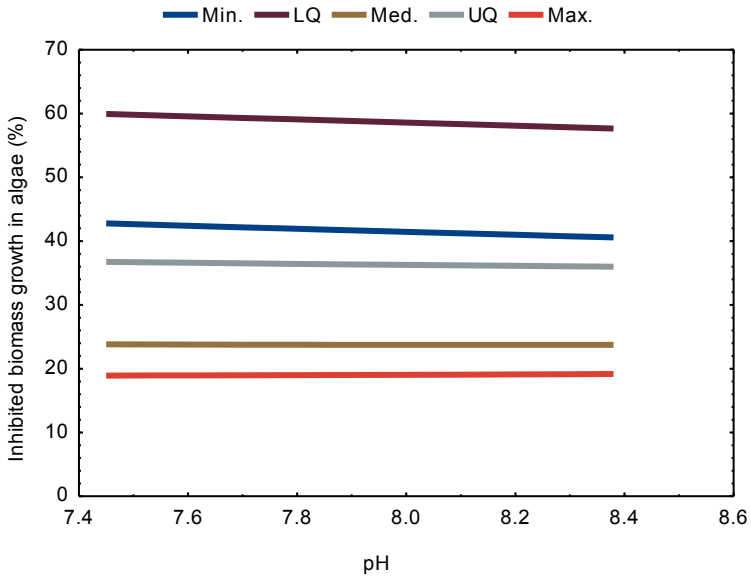


Figure 7.15 Influence of pH on inhibited biomass growth in algae.

In natural water, potential dormancy or death of algae species occurs in the pH ≥ 8.0 and ≤ 10.0 range (Pedersen and Hansen 2003), the region within which training data lies. The ENMs can trigger pH-induced inhibitions attributable to low CO_2 uptake and reduced metabolism in algae. However, the negligible model sensitivity to pH is a limitation of

available data to project response variants over a narrow range of values (pH > point of zero charge of $nTiO_2$) (Table 7.4) rather than zero contribution to toxicity that is initially high.

7.3.2 Results for Inhibited Reproduction in *D. magna*

Twenty-five models for each of the 8 network topologies were retained for evaluation. Similar to the previous models, the number of hidden neurons does not significantly influence the performance of models (Table 7.5) whose error variances are illustrated in Figures 7.16 and 7.17). The minimum validation set error was a criterion used to select and implement the ensemble model highlighted in grey (MLP 12-11-1) (Table 7.5). The prediction capability of the implemented model improved the performance of standalone models by 9.1%. The error of the full ensemble model based on the training and validation data sets was 0.396. This Section of the Chapter discusses the performance and sensitivity of the selected model.

Table 7.5 Summary of standalone and ensemble model performance

	Standalone model MSE (0.95 confidence interval)			Ensemble model MSE		Ensemble
Topology	Training	Validation	Testing	Testing	Reduction (%)	R^2
MLP 10-4-1	0.274 ± 0.027	0.290 ± 0.030	0.550 ± 0.048	0.472	14.2	0.600
MLP 10-5-1	0.260 ± 0.026	0.276 ± 0.033	0.547 ± 0.028	0.477	12.7	0.588
MLP 10-6-1	0.259 ± 0.024	0.271 ± 0.031	0.523 ± 0.028	0.466	10.8	0.590
MLP 10-7-1	0.241 ± 0.020	0.246 ± 0.022	0.513 ± 0.027	0.455	11.4	0.566
MLP 10-8-1	0.236 ± 0.017	0.247 ± 0.026	0.490 ± 0.022	0.443	9.4	0.570
MLP 10-9-1	0.227 ± 0.019	0.240 ± 0.021	0.488 ± 0.023	0.446	8.8	0.556
MLP 10-10-1	0.219 ± 0.012	0.233 ± 0.018	0.500 ± 0.020	0.460	8.2	0.542
MLP 10-11-1	0.216 ± 0.018	0.231 ± 0.021	0.498 ± 0.022	0.453	9.1	0.536

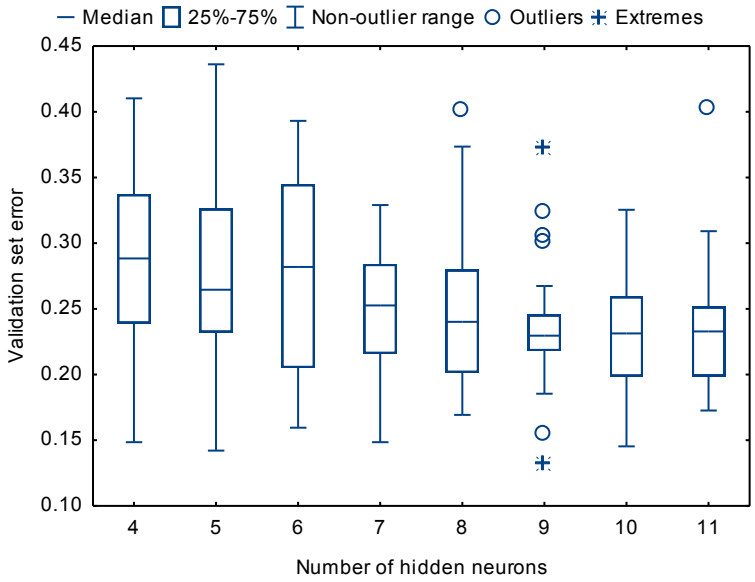


Figure 7.16 Average validation set errors of networks having a different number of hidden neurons.

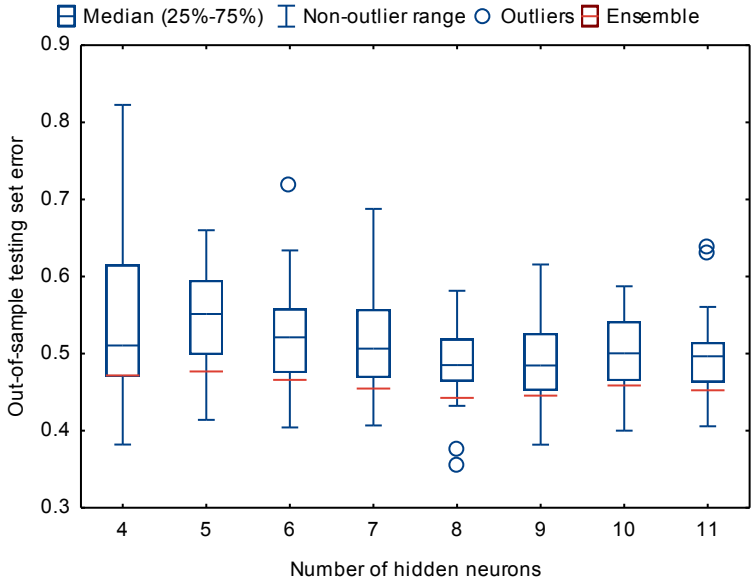


Figure 7.17 Testing set errors of networks having a different number of hidden neurons.

The ensemble model predicted known responses using withheld out-of-sample data. A scatter comparison of the prediction displays $R^2 = 0.534$ (Table 7.5 and Fig. 7.18), which sufficiently describes the inhibited reproduction in *D. magna*. The independence of predicted versus residual values ($R^2 = 0.001$ and slope of -4.73) shows no specific trend (Fig. 7.19) (*Left*). The residual values show under- and over-prediction of inhibited reproduction in *D. magna*, with an unbalanced right-skewed distribution having a normality of 0.934 (Fig. 7.19) (*Right)*.

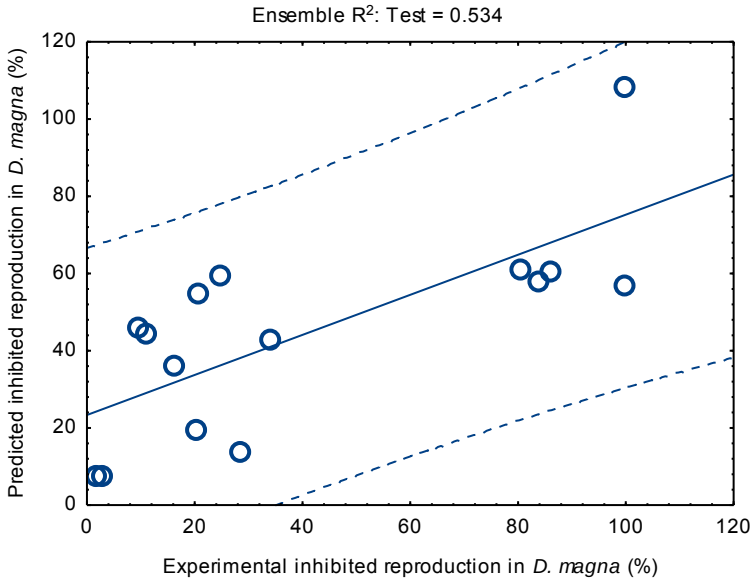


Figure 7.18 Ensemble prediction of inhibited reproduction in *D. magna.*

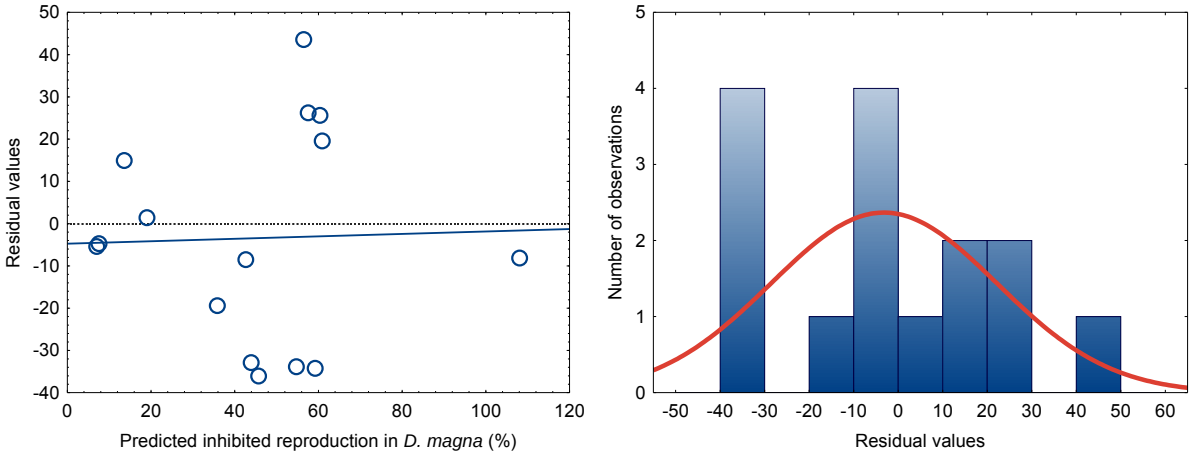


Figure 7.19 Predicted inhibited reproduction in *D. magna* versus residual values (*Left*) and normality of residual values (*Right*).

7.3.2.1 Relative Importance of Input Variables

Figure 7.20 presents the sensitivity (in bold) of input variables computed using perturbation MSE (in italics) against the MSE of the full model. The six most influential variables are organism generation > organism feeding > SAC > charged ions > ionic strength > initial total carbon content, an association comprising physical properties of $nTiO_2$ as well as and abiotic and biotic factors. In reverse order, the least important variables are pH, composition of anatase crystals, temperature, development stage, duration of exposure, and surface coating.

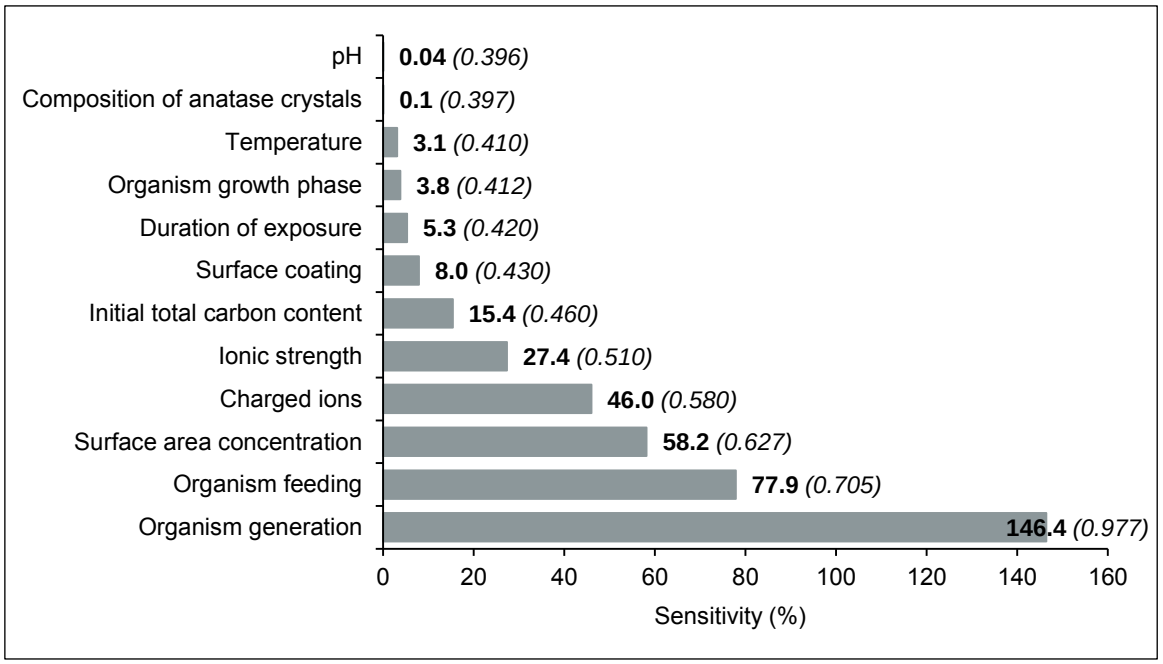


Figure 7.20 Sensitivity of input variables in predicting inhibited reproduction in *D. magna*.

Similar to the previous Section, the temperature and pH display low variable sensitivities in predicting the toxicity effects, conceivably explained by lack of significant variation in the data used to train the model. Independently, Table 7.6 presents the MSE of individually perturbed nominal level of categorical input variables and corresponding sensitivities (in italics).

In subsequent Sections (a) to (l), variable sensitivities presented in Figure 7.20 and Table 7.6 are discussed in tandem with model sensitivity to input perturbations using hypothetical data (Table 7.7). The input data of all variables display repeated values that limits the diversity of response trends estimated except for the SAC and duration of exposure (Table 7.7). The exposure conditions are basic (pH > point of zero charge of $nTiO_2$), . Whereas the total carbon content simulates large rivers and lake water, the ionic strength represents surface water (including rivers and lakes) (refer to Table 4.7 in Chapter 4).

The inhibitions on < 24-hrs neonates, < 24-hrs neonates, < 72-hrs neonates, < 24-hrs juveniles and < 24-hrs juveniles *D. magna* at the time of exposure to $nTiO_2$ are assigned $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ to identify outputs estimated using the minimum, LQ, median, UQ and maximum data, respectively.

Table 7.6 Sensitivity of nominal levels in predicting inhibited reproduction in *D. magna*

Input variable	MSE and sensitivity	Rank	Input variable	MSE and sensitivity	Rank
Organism growth phase			Organism generation		
Neonates < 24-hours	0.421 (*5.8*)	1	Sixth	1.133 (*65.0*)	1
Juveniles < 24-hours	0.414 (*4.3*)	2	Fifth	1.102 (*64.0*)	4
Neonates -72-hours	0.400 (*0.8*)	3	Fourth	1.072 (*63.0*)	5
Organism feeding			Third	1.040 (*61.9*)	4
0.3 mg C/ℓ	0.917 (*56.8*)	1	Second	1.010 (*60.7*)	5
1.5 mg C/ℓ	0.880 (*55.0*)	2	First	0.505 (*21.6*)	6
0.05-0.1 mg C/day (*Scenedesmus obliquus*)	0.798 (*50.3*)	3	Surface coating		
0.2 mg C/organism (*Desmodesmus sp.*)	0.629 (*37.0*)	4	AlPDMSGl	0.443 (10.6)	1
1 x 10^6 cells/mℓ /day	0.533 (*25.6*)	5	AlPDMS	0.442 (10.4)	2
YCT and 1 x 10^7 cells/mℓ/day	0.473 (*16.3*)	6	Bare	0.398 (0.4)	3
Charged ions			-	-	-
Elendt 4	0.694 (*42.9*)	1	-	-	-
RDCM	0.678 (*41.6*)	2	-	-	-
MHW	0.476 (*16.8*)	3	-	-	-
HW	0.467 (*15.2*)	4	-	-	-

Table 7.7 Hypothetical exposure conditions to project inhibited reproduction in *D. magna*

Input variable	Minimum	LQ	Median	UQ	Maximum
SAC (m^2/ℓ)	3.00E-3	0.06	0.10	0.24	0.60
Temperature (°C)	19.85	21.35	25.00	25.00	25.00
Initial carbon content (mg/ℓ)	8.10	9.37	13.73	13.73	27.46
Duration of exposure (hours)	72.0	192.0	336.0	384.0	504.0
Organism feeding	YCT and 1E7 algae cells/day	YCT and 1E7 algae cells/day	YCT and 1E7 algae cells/day	0.2 mg C/organism /day	0.3 mg C/ℓ
Charged ions	MHW	MHW	MHW	HW	Elendt 4
Surface coating	Bare	Bare	Bare	Bare	AlPDMSGl
Composition of anatase crystals (%)	0.0	70.0	70.0	80.0	99.0
pH	7.60	7.60	7.60	8.00	8.26
Ionic strength (mM)	4.24	4.24	4.24	8.23	8.51
Organism generation	First	First	First	First	Sixth
Organism growth phase	Neonates (< 24-hrs)	Neonates (72-hrs)	Neonates (72-hrs)	Juveniles (< 24-hrs)	Juveniles (< 24-hrs)

a) Organism Generation

The high sensitivity of generational effects of $nTiO_2$ on *D. magna* indicates the significance of this variable in nanoecotoxicology assessments (Fig. 7.20). There is a similarity in the ranking of individual sensitivities (Table 7.6 and Fig. 7.21), confirming findings by Jacobasch et al. (2014) that exhibited a population collapse in the sixth generation. Consequent on exposure conditions, the toxicity effects on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ in estimated in the first generation increased by 70.8%, 60.9%, 51.3%, 99.7%, and 78.0% in the sixth generation, respectively (Fig. 7.21). Thus, a rise in generational inhibitions as a manifestation of toxicity effects in the offspring of surviving *D. magna*.

In all generations, the magnitude of inhibitions on $N72_{(med.)} > N24_{(LQ)} > N24_{(m)} > J24_{(max)} > J24_{(UQ)}$ suggest higher toxicity on *D. magna* exposed at neonate than juvenile stage (also discussed in Section (i)), which confirms published literature (Wright and Welbourn 2009). Thus, acute and chronic ecotoxicology studies may underestimate generational impacts of ENMs on biological organisms relevant to the aquatic environment. The relevance of this

variable in modelling is the potential to establish inhibitions to organism populations that can disrupt functional mechanisms, for example, the food chain in aquatic systems.

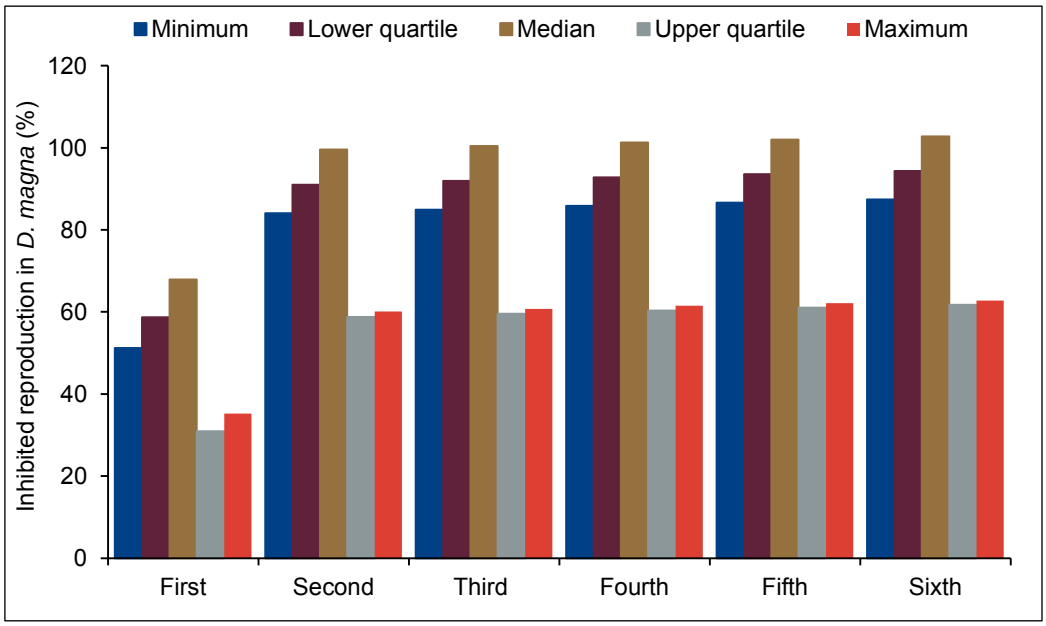


Figure 7.21 Influence of organism generation on inhibited reproduction in *D. magna*.

b) Organism Feeding

The ranking of individual feed types presented in Table 7.6 is inverted (Figure 7.22). The model appears to be greatly sensitive to low ranked feed types and vice versa. Contingent on simulated exposure conditions, the inhibitions on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$ and $J24_{(max)}$ decrease by 62.3%, 54.4%, 48.0%, 53.0% and 30.8%, respectively, in a consistent sequence of 6 > 5 > 4 > 3 > 2 > 1 (refer to Table 7.6 to identify feed types) (Fig. 7.22). Besides providing nourishment to *D. magna*, the effects of food on $nTiO_2$ stability (Campos et al. 2013) and detoxification (Zhu et al. 2010) govern bioavailability and toxicity outcomes.

The diverse proportions of algae species fed to *D. magna* create a pseudo-mixed culture exposure system that clearly shows disparate influence on toxicity but information on their composition is insufficient for meaningful comparisons at this stage. However, the model has demonstrated the usefulness of this variable in explaining toxicity and the need for full characterization in exposure studies to support nanoecotoxicology research.

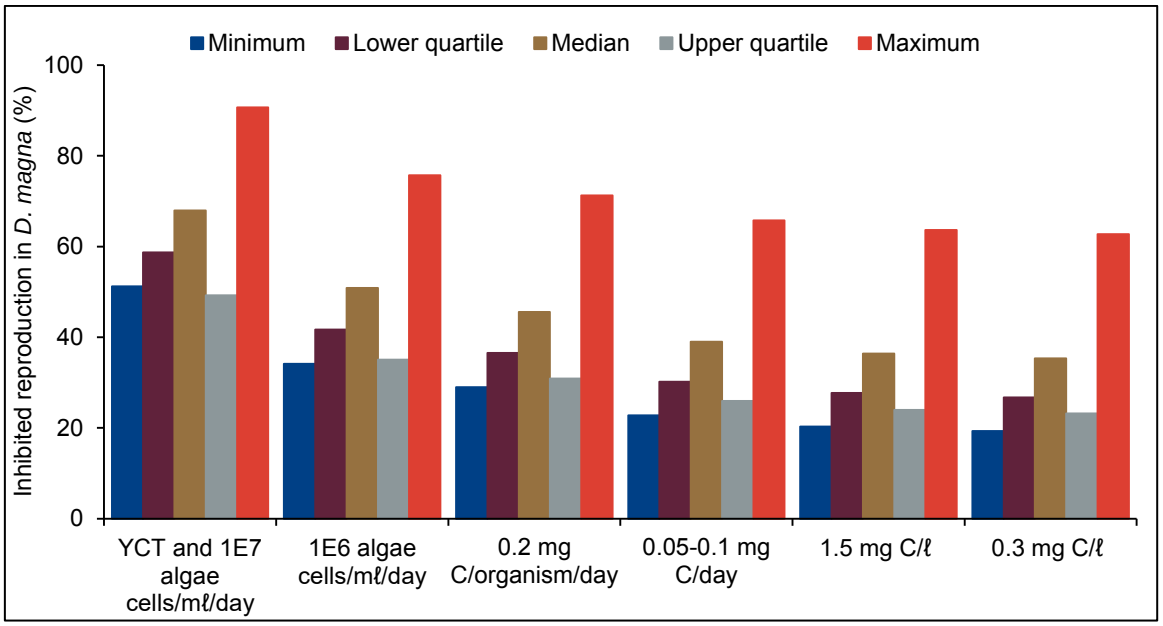


Figure 7.22 Influence of organism feeding on inhibited reproduction in *D. magna*.

c) Surface Area Concentration

With a sensitivity of 58.2%, the SAC displays very low values in the minimum-maximum range (Table 7.7) compared to those used to train previous models that suggest higher physical-related than chemical-related toxicity interactions of $nTiO_2$ with *D. magna*. The inhibitions to $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ at 0.003 m^2/ℓ increase by 122.3%, 117.9%, 104.1%, 517.1%, and 290.8% at 0.6 m^2/ℓ, respectively (Fig. 7.23).

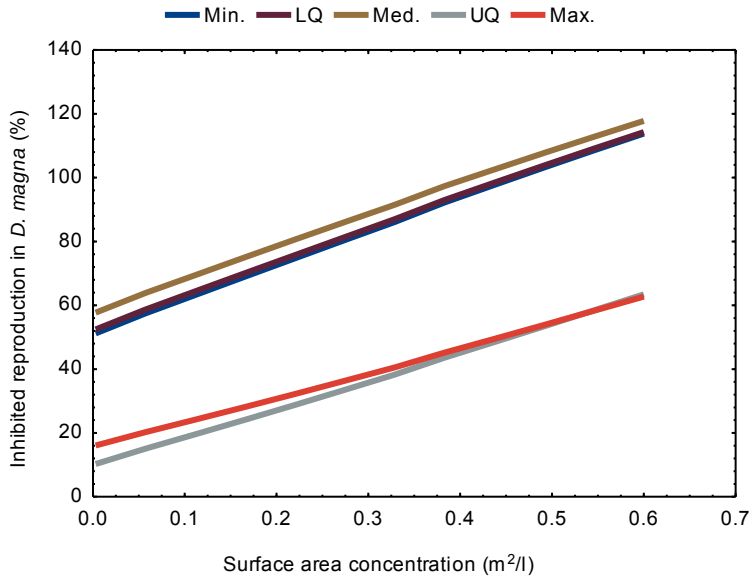


Figure 7.23 Influence of surface area concentration on inhibited reproduction in *D. magna*.

The toxicity of $nTiO_2$ to *D. magna* has been linked to decreasing primary size and increasing concentration of the ENM (Warheit et al. 2007; Kim et al. 2010; Zhu et al. 2010; Dabrunz et al. 2011; Das et al. 2013; Clément et al. 2013; Jacobasch et al. 2014). The findings do not directly relate to effects of SAC but the trends clearly show enhanced toxicity with increasing surface-to-volume ratio of the ENM. Moreover, an *ca.* 1:95 increase in SAC:inhibition shows the necessity of utilizing this variable to evaluate risk profiles of ENMs.

d) Charged Ions

The projected inhibitions in RDCM, HW, and MHW test mediums increase by an average of 1.01-fold, 1.24-fold, and 1.47-fold in Elendt 4 medium, respectively (Fig. 7.24) under all exposure conditions, which is a similar ranking presented in Table 7.6. In the order: MHW < HW < RDCM < Elendt 4, the toxicity of $nTiO_2$ on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ increases by 45.0%, 38.7%, 32.8%, 76.7%, and 41.6%, respectively (Fig. 7.24).

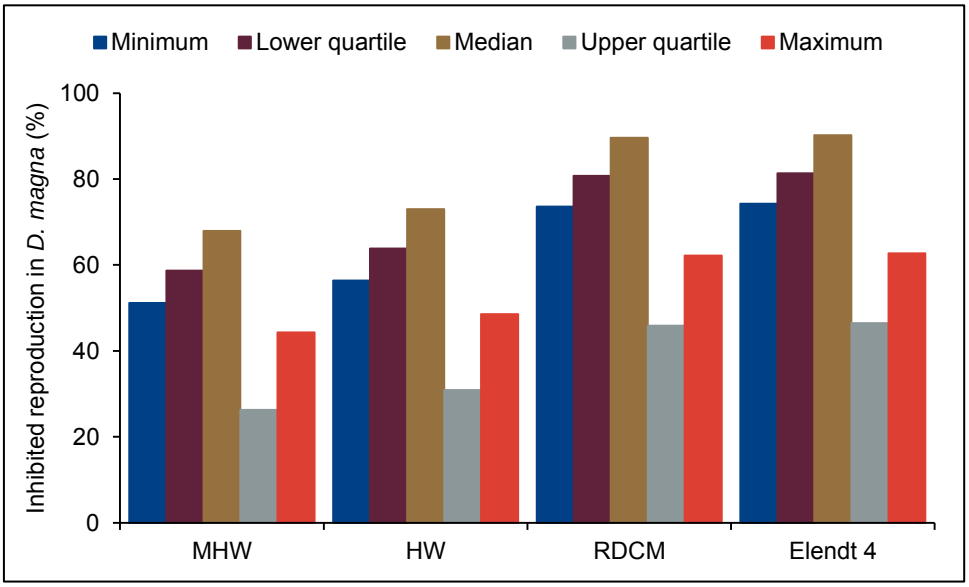


Figure 7.24 Influence of charged ions on inhibited reproduction in *D. magna*.

The concentration of Ca^+ ions in Elendt 7 test medium, which is closer to Elendt 4, was found to influence the uptake and accumulation of dispersed or aggregated $nTiO_2$ in *D. magna* (Tan et al. 2016). The authors (Tan et al. 2016) focused on evaluating Ca^+ ions only to simplify analyses but the model results depict an increase in inhibitions based on diverse concentration and speciation of mixed ions in the exposure media and resultant dispersity of ENMs. The $nTiO_2$ is positively charged in Elendt 7 (Cupi et al. 2015), which illustrates higher interactions with negatively charged *D. magna* in Elendt 4 than in MHW and HW where the ENM is negatively charged (Ottofuelling et al. 2011). Thus, the richness of micro- and

macro-nutrients in Elendt 4 followed by the RDCM, and basic macronutrients in the HW and MHW, explain diverse contribution of mixed ions on toxic inhibitions. However, the model contrasts findings (Seo et al. 2014) of higher toxicity in MHW than HW attributed to enhanced aggregation in the latter medium.

e) Ionic Strength

A 4.28 mM increase in ionic strength decreases the toxicity effects of $nTiO_2$ on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ by 62.9%, 55.0%, 48.7%, 49.1%, and 31.6%, respectively (Fig. 7.25). The previous Chapters reveal a marginal increase in organic adsorbates on $nTiO_2$ (Fig. 5.12) and an increase or decrease in aggregation (Fig. 6.7) with increasing ionic strength. Thus, ionic strength indirectly influences toxicity by way of induced stability and bioavailability of $nTiO_2$, which in this case shows an interaction barrier between *D. magna* and the ENMs that reduces potential toxic effects.

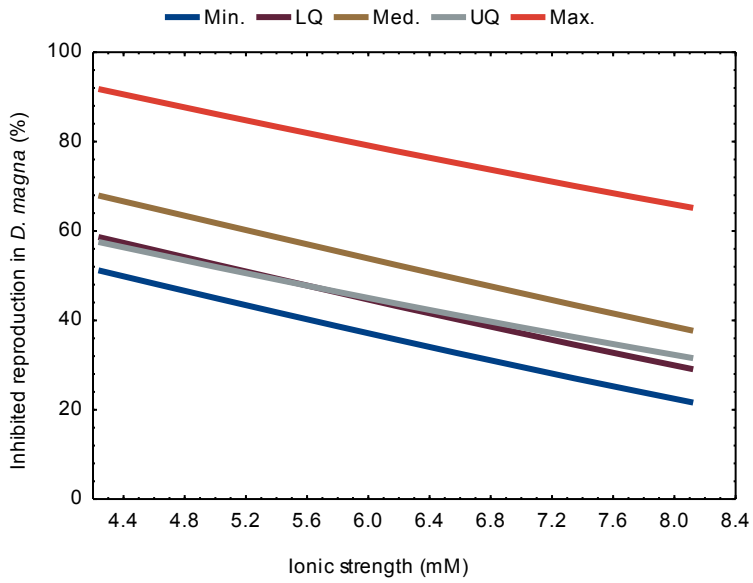


Figure 7.25 Influence of the ionic strength on inhibited reproduction in *D. magna*.

f) Initial Total Carbon Content

The toxicity effects of $nTiO_2$ on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ decrease by 77.6%, 65.0%, 50.0%, 80.5%, and 33.7% with increasing initial total carbon content (Fig. 7.26), a trend that confirm published literature (Cupi et al. 2015). The explanatory contribution of this variable to rising adsorption of organic matter on $nTiO_2$ (Fig. 5.9), rising or reducing aggregation of the ENM (Fig. 6.15) and decreasing toxicity to algae feed (Fig. 7.10) hint at reduced bioavailability, improved depuration and decreased short-term toxicity

effects. Thus, it is likely that under environmental conditions, the levels of carbon can mitigate the negative impacts of ENMs on biological organisms.

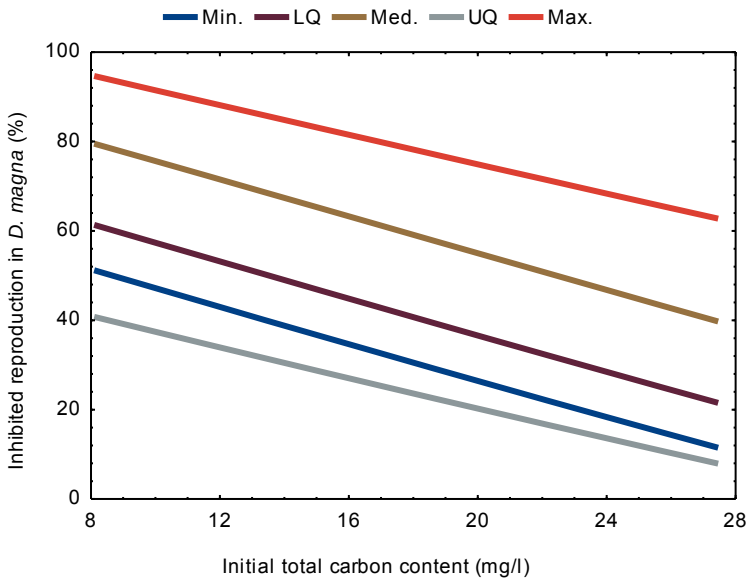


Figure 7.26 Influence of initial carbon content on inhibited reproduction in *D. magna*.

g) Surface Coating

The projected toxicity of bare $nTiO_2$ on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ reduces by *ca.* 1.21-fold, 1.18-fold, 1.15-fold, 1.30-fold, and 1.12-fold because of capping with AlPDMSGl and AlPDMS (7.27), which inverts the ranking order presented in Table 7.6.

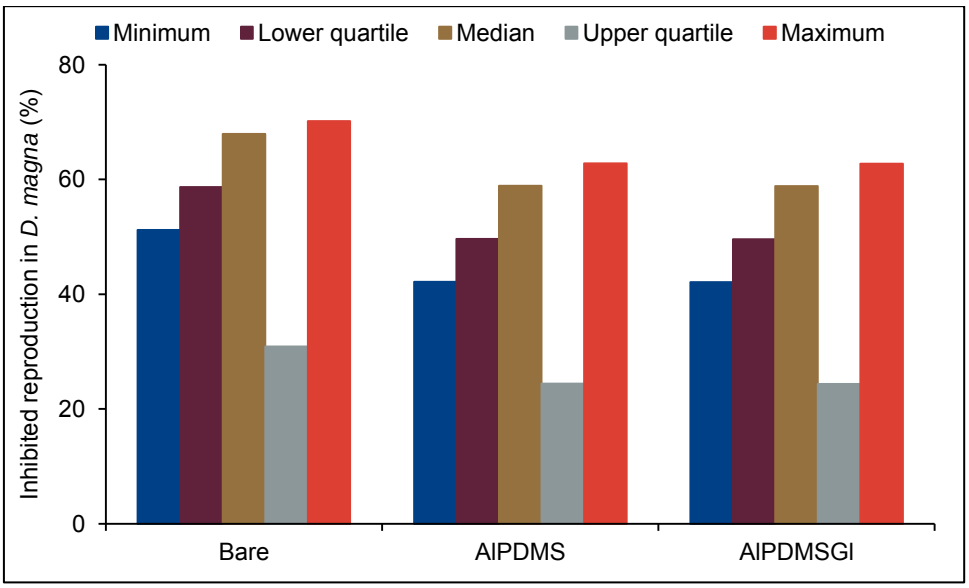


Figure 7.27 Influence of ENMs' surface coating on inhibited reproduction in *D. magna*.

The model findings agree with remarks by Das et al. (2013) implying toxicity alleviation by capping agents. The toxicity effects induced by capped $nTiO_2$ is slightly lower than that from bare nanoparticles (Fig. 7.27), which suggest reduced bioavailability because of settled solids or transformed surface charges based on results presented in Figure 6.5. Thus, both capped and uncapped ENMs induce toxicity but the degree of effects depends on exposure conditions. Thus, information about properties of capping agents is useful in nanoecotoxicology research to estimate likely direct and indirect risks of ENMs.

h) Duration of Exposure

Under all exposure scenarios, an increment of 39.3-hours results in 0.8-1.0% increase in toxicity (Fig. 7.28). The projections imply marginal toxicity effects on *D. magna* after lengthy exposure periods to $nTiO_2$. The inhibitions on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ at 72-hrs increase by 20.2%, 18.3%, 16.2%, 52.9%, and 16.2% at 504-hours, respectively, a trend that agrees with findings (Clément et al. 2013; Kim et al. 2014; Seitz et al. 2015). Thus, data produced under lengthy exposures would yield superior model estimation of inhibitions than shorter exposures, considering generational effects presented in Section (i) above.

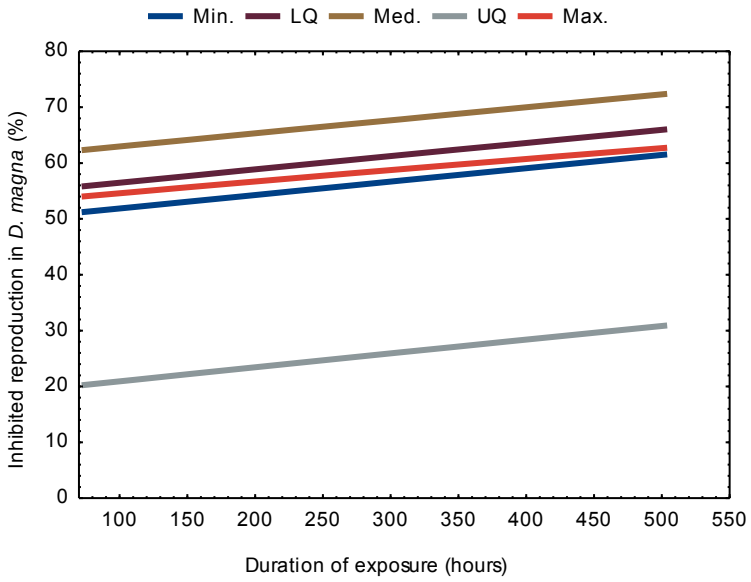


Figure 7.28 Influence of duration of exposure on inhibited reproduction in *D. magna*.

i) Organism Growth Phase

The neonate (< 24-hours) > juvenile (< 24-hours) > neonate (72-hours) sensitivity (Table 7.6) and projected inhibitions (Fig. 7.29) suggest slightly high vulnerability of < 24-hours neonate *D. magna* to ENMs toxicity. The model findings agree with remarks by Rand et al. (2003)

that linked poor pollutant detoxification mechanisms of neonatal stage organisms to enhanced toxicity effects. The reversed ranking of < 72-hrs neonates and < 24-hrs juvenile is a pointer to model sensitivity uncertainties because of data used or diverse susceptibility dictated by exposure conditions. As an indicator of pollutant susceptibility (refer to Section 4.6.4 in Chapter 4) and population growth of surviving organisms (Eqn. 7.1), the initial organism growth phase is a variable worth exploring in predictive modelling to estimate ecological effects of ENMs.

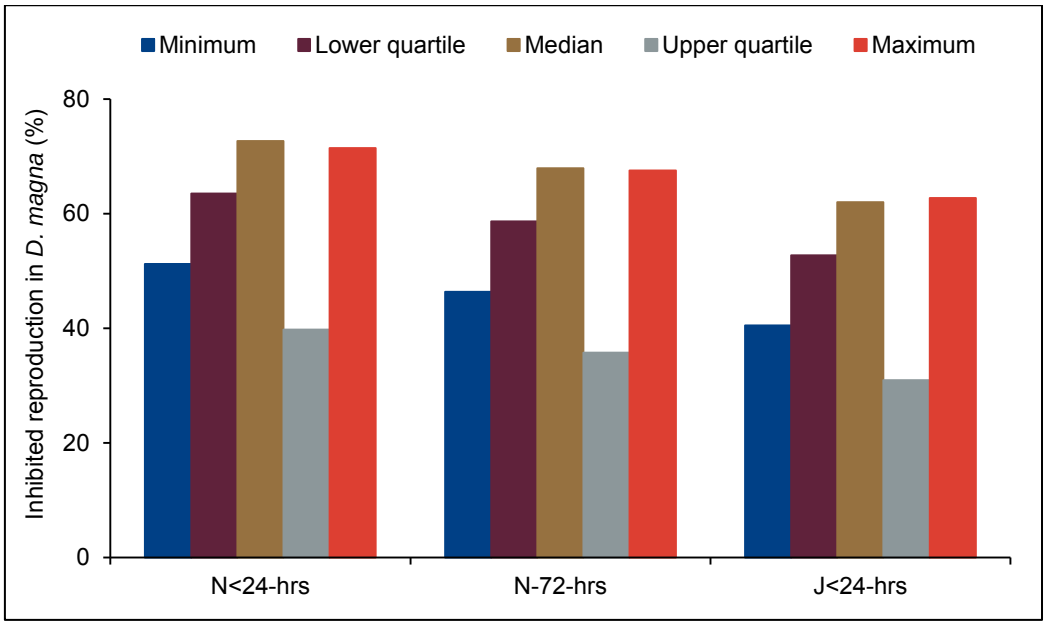


Figure 7.29 Influence of organism growth phase on inhibited reproduction in *D. magna*.

j) Temperature

The sensitivity of temperature is only 3.14% (Fig. 7.20) but its contribution to toxicity seems to be essential when a 0.47 ^{o}C change results in a 1.9-fold to 3.2-fold (no order) increment in inhibited reproduction. The inhibitions on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ reproduction increased by 20.1%, 26.4%, 27.4%, 48.0%, and 21.0% with increasing temperature, respectively (Fig. 7.30). The previous model projections showed decreased adsorption of organic matter on $nTiO_2$ (Fig. 5.4), increased or decreased aggregation of $nTiO_2$ (Fig. 6.6) and near constant but substantive effects of $nTiO_2$ on algae (Fig. 7.14) with increasing temperature. Therefore, the temperature indirectly influences aggregation and settlement of algae-$nTiO_2$, which has been shown to deplete food needed by *D. magna* (Campos et al. 2013) and hinder feeding mechanisms (Kim et al. 2010; Zhu et al. 2010; Jacobasch et al. 2014) that reduces survival of the organism.

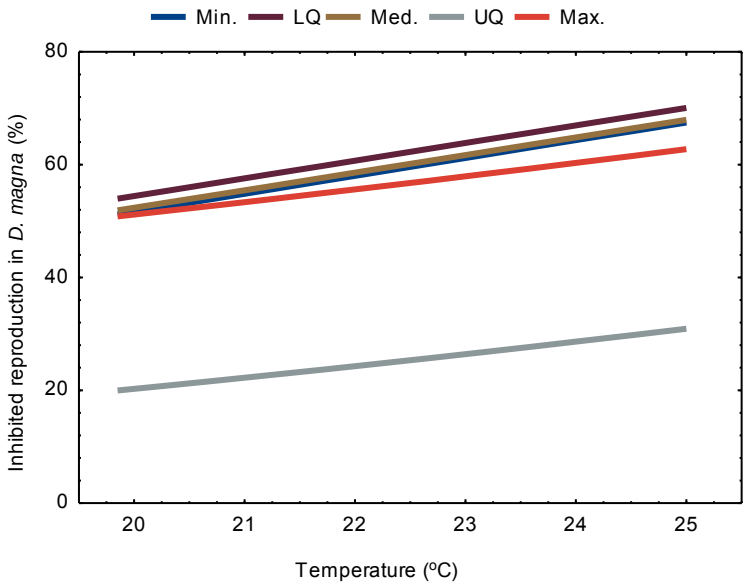


Figure 7.30 Influence of the temperature on inhibited reproduction in *D. magna*.

Moreover, an increase in temperature can directly decrease the solubility of oxygen (Wright and Welbourn 2009), which depletes the resource (O_2) needed by phytoplankton, a condition that enhances metabolism in *D. magna* as a defense mechanism (Khan and Khan 2008). However, an increase in metabolism requires extra but non-available food and oxygen and also leads to hypoxia that reduces size and mass of *D. magna* (Khan and Khan 2008), thus, reproduction capability (Eqn. 7.1). Thus, a synergy between the two toxicity scenarios requires additional investigations.

k) Composition of Anatase Crystals

The model sensitivity to anatase crystals in elucidating inhibited reproduction in *D. magna* (Fig. 7.31) reveals the implication of the low sensitivity (0.10%) presented in Figure 7.20. Despite a 29.0% to 66.9% initial inhibition (no order), a simulated 9.0% increase in anatase $nTiO_2$ crystals effected a 0.01-0.02-fold increment in toxicity inhibitions. The initial inhibitions on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ increased by 3.1%, 2.7%, 2.1%, 8.2%, and 3.3%, respectively. Hence, the low sensitivity is a hint of a small margin between toxicity effects induced by anatase, anatase / rutile mixture, and rutile $nTiO_2$ crystals. The trends correspond to findings (Clément et al. 2013;Seitz et al. 2013) that indicated higher toxicity effects of anatase than anatase / rutile mixture or rutile, attributable to different crystallinity and reactivity discussed in Section 4.4.1 in Chapter 4.

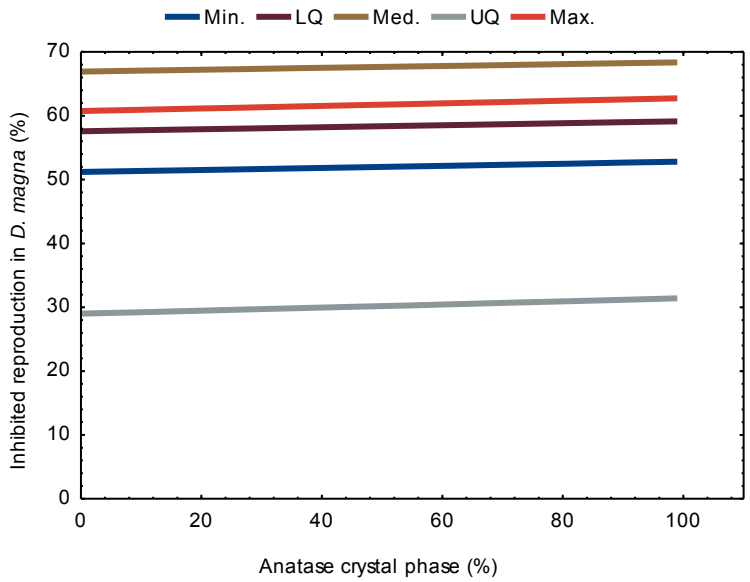


Figure 7.31 Influence of the composition of anatase crystals on inhibited reproduction in *D. magna*.

l) pH

The pH displays the least sensitivity (0.04%) to inhibited reproduction in *D. magna* (Fig. 7.20), which may also be attributed to the narrow range of values used to train the model (Table 7.7). However, initial high inhibitions (36.0% to 73.0%, no order) and a 0.06 change in pH that induces a 1.7-fold to 23.4-fold decrease in the inhibitions, reveals a considerable contribution of the variable to predicting toxicity effects of $nTiO_2$ on the organism.

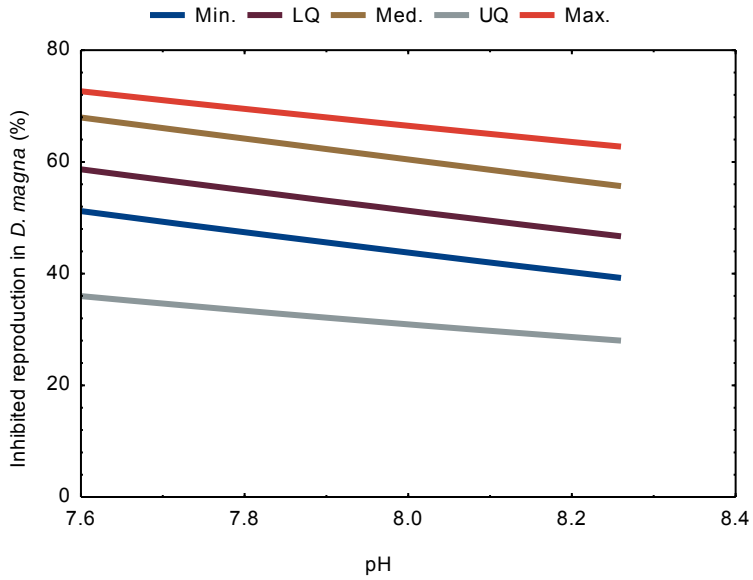


Figure 7.32 Influence of the pH on inhibited reproduction in *D. magna*.

The inhibitions on $N24_{(m)}$, $N24_{(LQ)}$, $N72_{(med.)}$, $J24_{(UQ)}$, and $J24_{(max)}$ decrease by 18.1%, 18.2%, 18.6%, 12.1%, and 15.0%, respectively, with increasing pH. However, it is challenging to relate the trends with previous model projection that showed a decrease in organic adsorbates on $nTiO_2$ (Fig. 5.10), increase or decrease in aggregation (Fig. 6.11) and slight decrease or increase toxicity to algae (Fig. 7.13) with increasing pH.

7.4 Projected Associations amongst Adsorption, Aggregation and Toxicity Effects

This Section of the Chapter discusses neural networks' model simulations of exposure scenarios that facilitate the understanding of associations between the behaviour of $nTiO_2$ and biological toxicity on algae and *D. magna* in aqueous environment. The application of the model discussed in Chapters 5 and 6 to fit appropriate data used to train models presented in this Chapter aimed to achieve the following

- Predict the quantity of organic adsorbates on $nTiO_2$ and hydrodynamic size of the ENM.
- Utilize the estimated outputs of behaviour response outputs to establish likely relationships with toxicity of $nTiO_2$ presented in this Chapter using scatter plots.

To a certain extent, the uncorroborated forecasts of adsorbed mass and hydrodynamic size test the prediction capability of models discussed in Chapters 5 and 6 using new data.

7.4.1 Estimates using Data for Inhibited Biomass Growth in Algae

The relationship between hydrodynamic size of $nTiO_2$ and inhibited biomass growth in algae is nonlinear and without a specific pattern (Fig. 7.33). Notably, an increase in $nTiO_2$ size induces a high percentage of toxicity on algae species. In contrast, a decrease in $nTiO_2$ size contributes a small degree of low and high toxicity effects. Moreover, Figure 7.34 demonstrates a nonlinear scatter comparison between adsorption and aggregation behaviour of $nTiO_2$ and toxicity of the ENM on algae. The following describes the associations depicted in Figure 7.34,

- A decrease or minimal increase in organic adsorbates on $nTiO_2$ and high dispersity of increasing aggregate sizes results in ≥ 60.0% inhibitions of biomass growth in algae.
- A decrease or increase in organic adsorbates on $nTiO_2$ and low dispersity of decreasing aggregate sizes results in ≤ 60.0% inhibitions of biomass growth in algae.

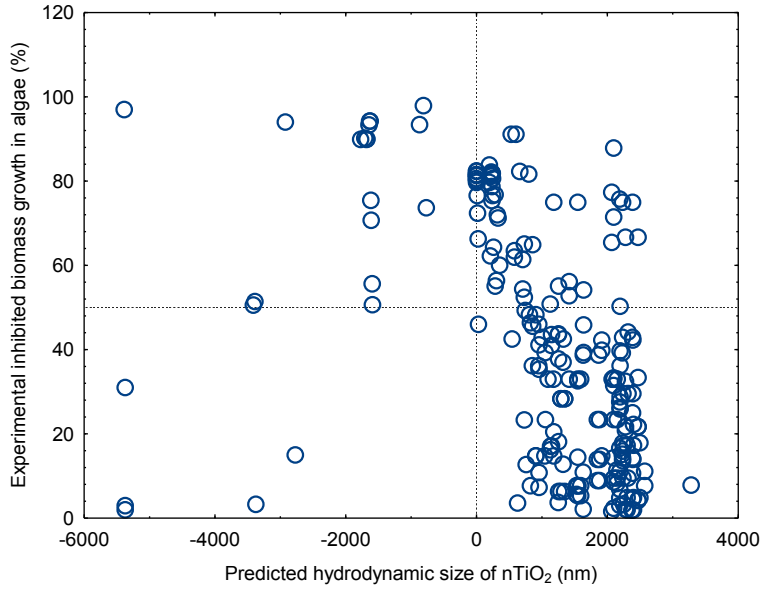


Figure 7.33 Predicted hydrodynamic size versus experimental inhibited biomass growth in algae.

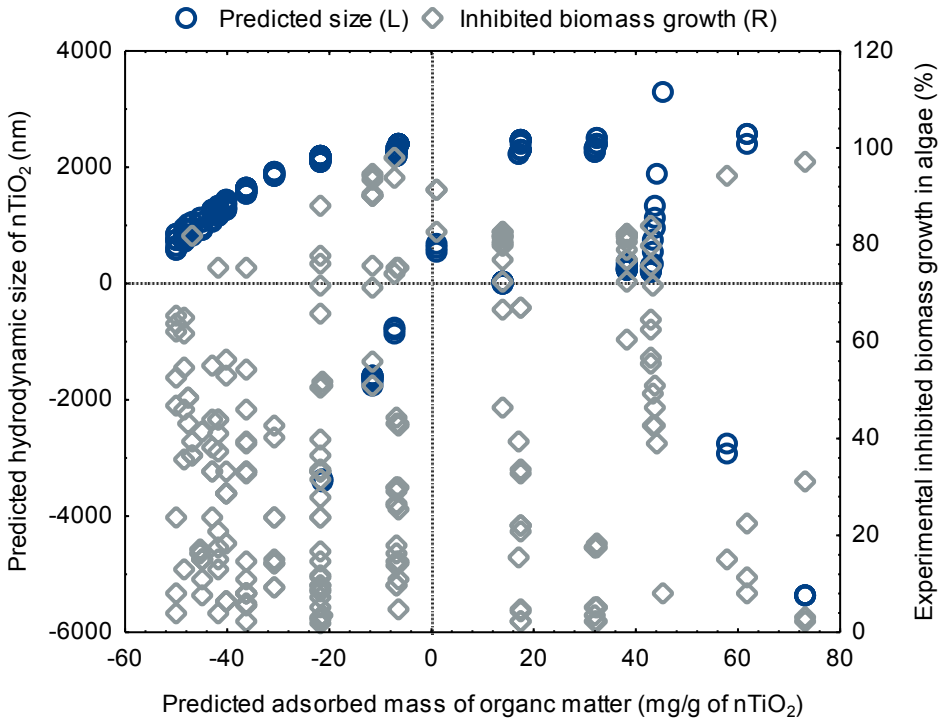


Figure 7.34 Association between adsorbed mass, hydrodynamic size, and inhibited biomass growth in algae.

Thus, the MLP-NN model has competently extracted patterns from the data that agrees with findings in published literature (Aruoja et al. 2009; Hartmann et al. 2010; Sadiq et al. 2011; Ji et al. 2011; Metzler et al. 2011), which show $nTiO_2$ aggregates encapsulating algae species.

The results show potential formation of diverse fractals of ligands-ENMs at different levels of bioavailability, accumulation, and trophic transfer in aquatic organisms.

7.4.2 Estimates using Data for Inhibited Reproduction in *D. magna*

The association between hydrodynamic size of $nTiO_2$ and inhibited reproduction in *D. magna* is nonlinear and nonspecific (Fig. 7.35). The toxicity effects are high when the decrease in aggregate size (likely small particles / aggregate solids of $nTiO_2$) is high and vice versa.

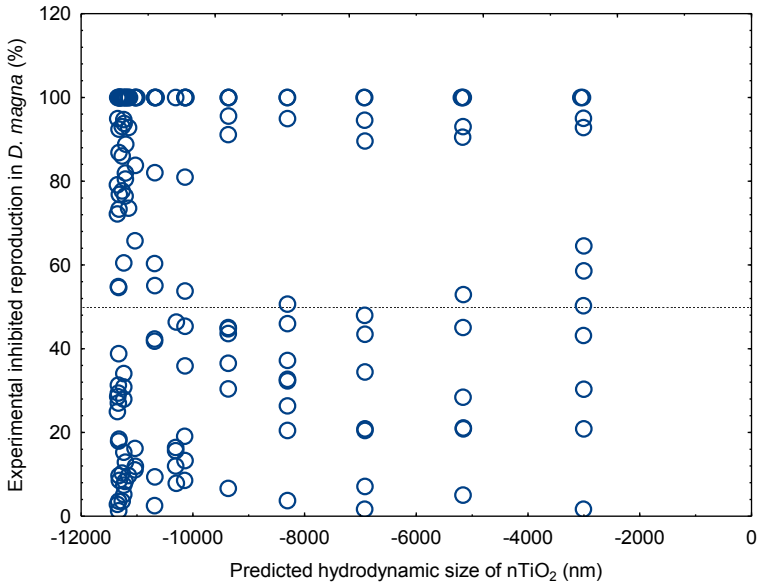


Figure 7.35 Predicted hydrodynamic size versus inhibited reproduction in *D. magna*.

Moreover, Figure 7.36 also illustrates a nonlinear association between adsorption and aggregation behaviour and toxicity of $nTiO_2$ on *D. magna* described thus,

- A decrease in aggregation and adsorption contributes to a large percentage of low and high inhibitions of $nTiO_2$ on *D. magna* without a specific trend.
- The inhibitions are ≤ 50.0% under low adsorption of organic matter on $nTiO_2$ in the absence of aggregation, ≥ 50.0% under similar conditions with marginal reduced aggregation.

Similarly, the MLP-NN model has competently extracted the relationship between the behaviour of $nTiO_2$ and toxicity on *D. magna*. However, unlike the association observed in Figures 7.33 and 7.34, toxicity of $nTiO_2$ on *D. magna* appears influenced by agglomerates (particle-particle interactions) or dispersed solids of the ENM.

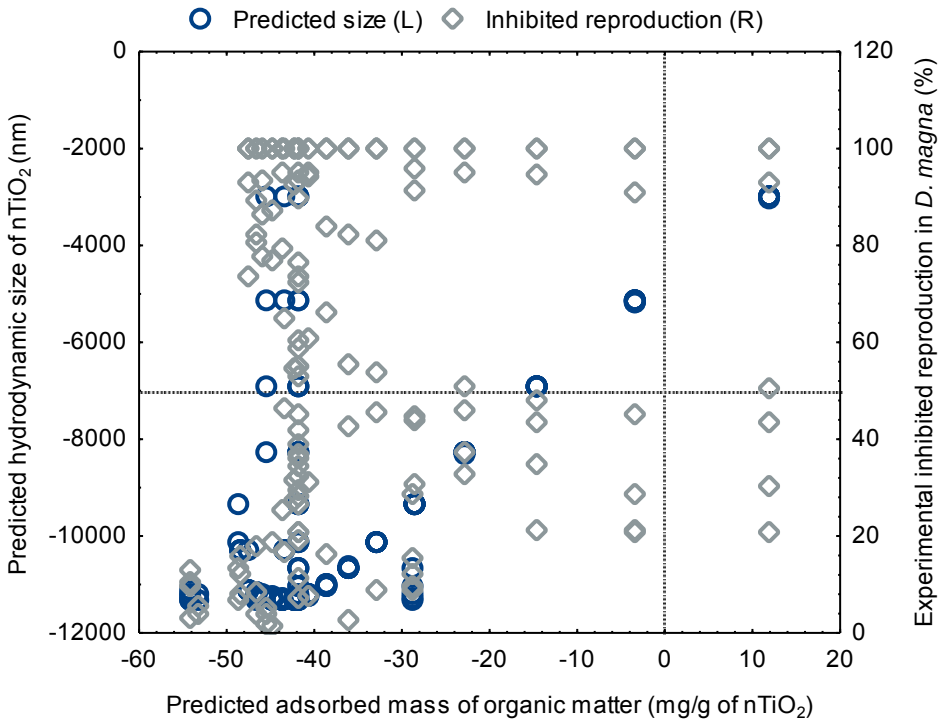


Figure 7.36 Association between adsorbed mass, hydrodynamic size, and inhibited reproduction in *D. magna*.

The high and low SAC values recorded in the algae dataset (Table 7.4) and *D. magna* dataset (Table 7.7), respectively, may explain the disparity in the projected exposure and toxicity scenarios (Figs. 7.34 and 7.36). Thus, there is likely differences in the surface charge, interaction, and toxicity of $nTiO_2$ having smaller surface area-to-volume-ratios than larger ones. Moreover, the differences in abiotic and biotic factors in the exposure media uniquely affect the $nTiO_2$-organism interactions (Sections 7.4.1 and 7.4.2), especially the magnitude and sign of surface charges considering negatively charged organisms also attract cations. Thus, the model has extracted adsorption-aggregation-toxicity relationships that reveal the influence of ENMs behaviour on the degree of toxicity in organism hinged on properties of exposure environment.

7.5 Conclusion

The MLP-NN model competently predicted linear inhibitions on biomass growth in algae and reproduction in *D. magna* using qualitative and quantitative data. The model appears to be a relevant tool for application in nanoecotoxicology research to predict similar occurrences in natural water when relevant data is available. The abiotic and biotic factors that promote functional mechanisms of organisms in simulated environmental habitats have varying

degrees of influence on $nTiO_2$ behaviour and toxicity. The model reveals the heterogeneous organism species and community in habitats and surviving offspring as highly ranked biotic factors that simulate two levels of the food chain likely to elevate toxicity effects.

The model findings demonstrate reduced toxicity levels of capped $nTiO_2$ on the organisms, which is a pointer to using capping agents as predictors in the absence of quantitative surface charge values as well as safe production and application of ENMs in nanoproducts. The contribution of SAC to predicting toxicity effects based on low values implies low concentration or small particles of $nTiO_2$ in an aquatic environment would be potentially toxic on biological organisms.

The model has demonstrated that mixed ions in exposure media affect the magnitude and sign of $nTiO_2$ surface charges. The exposure conditions that promote negative surface charges of the ENM would reduce toxicity and vice versa because of complexation of aggregation or dispersion scenarios. Thus, the application of toxicity data generated in diverse synthetic media limits inference of findings in natural water. However, the model's capability to predict toxicity responses using charged ions as a qualitative predictor shows likely deconstruction of exposure media diversities to create quantitative predictors applicable in projecting outputs relevant to an aquatic system.

The nonlinear relationships between adsorption, aggregation, and toxicity responses illustrate specific behaviour of ENMs and toxicity effects in different aquatic systems. Thus, it is likely that stability of ENMs in an environment that is rich in organic and inorganic matter would promote high but short-term effects. Conversely, aggregation and settlement of ENMs in an environment that is poor in organic and inorganic matter may stimulate low short-term effects. However, intermediate- and long-term generational effects in both systems are likely considering the simulated small increases in toxicity with increasing duration of exposure.

The significance of the MLP-NN models in an aquatic environment lies in the potential application of multiple interacting factors to estimate the effects of ENMs on organism population based on inhibited biomass growth in algae and reproduction in *D. magna*. The associated toxicity risks would also inform about water quality issues. The inclusion of environmental factors standardized in test protocols as model predictors has revealed a potential synergy between direct and indirect effects of these factors on toxicity not yet investigated. In spite of despite their contribution to high initial toxicity effects, the sensitivity of the two models to temperature and pH was very low; thus, the use of a broad range of values would support modelling research.

Chapter 8 - Summary, Conclusions, and Recommendations

This Chapter summarizes the thesis by revisiting the research questions and objectives presented in Chapter 1 to appraise achievements and limitations of this study. Conclusions are drawn from the findings and areas requiring further research recommended based on the limitations

8.1 Summary

This thesis discusses the application of multilayer perceptron (MLP) neural networks model as a data-driven computational technique to learn from scientific data in nanoecotoxicology domain and extraction of useful information. The technique, applied for the first time, approximates responses identified from behaviour kinetics and toxicity of ENMs in an aqueous environment. While the nature of data used in this study determined the batch mode modelling approach, the MLP model is adaptable and applicable in forecasting dynamic endpoints in aquatic environment based on continuously changing inputs.

Chapter 1 introduces engineered nanomaterials (ENMs) and nanostructures, their release into aquatic environment, and environmental concerns under nanoecotoxicology. Physicochemical properties of ENMs found to be suitable for industrial application raises concerns about their persistence and impacts in the environment. Review of five notable metal, carbonaceous and metal oxide ENMs and their release pathways identify nAg, $nTiO_2$, and nZnO as major aquatic pollutants. However, high estimated quantities produced and released into freshwater necessitated the selection of $nTiO_2$ as a model ENM to assess the kinetics and toxicity effects. The complex interactions of abiotic and biotic factors determine the health of aquatic environment, and surface transformation, fate, bioavailability and toxicity of discharged ENMs that may affect normal core functions in the longer-term.

The challenges arising from existing nanoecotoxicology strategies form the basis of this study. Chapter 1 discusses the limitations of conducting environmental risk assessment of ENMs, which motivated the application of computational data-driven modelling founded on the following reasons

- The use of diverse ENMs and nanostructures in nanotechnology products is uncertain and their risk indicators are not well defined.

- Low levels of accountability in nanotechnology production value chain, for instance, voluntary content and quantity labelling of ENMs in nanoproducts, affects the capacity to monitor risks.
- Unknown quantities and properties of nanowaste released into the environment.
- Involuntary risks posed by nanowaste in the absence of universal regulation and legislation on mitigation, monitoring, risk assessment and management of ENMs in the environment.

A set of research questions and objectives useful for data learning developed guided the study. Chapter 2 reviewed historical and current perspective of nanotechnology, production, application and release of ENMs into the aquatic environment. Also discussed are principles governing particle kinetics that provide scientific background on charged solids like ENMs in the aquatic environment. Chapter 2 discusses modelling approaches previously applied in nanoecotoxicology, whose strengths and weaknesses helped to identify input variables and construct a research methodology for this study. Chapter 3 presents neural networks technique, a component of knowledge discovery from data (KDD) methods presented in Chapter 2, and the research method applied in this study founded on ensemble learning paradigm.

Discussed in Chapter 4 are processes for compiling data for learning, outlining the framework and criteria used to screen and extract quantitative and qualitative data from the accessible scientific literature, and develop a manual registry. From the registry, a sample dataset was constructed comprising, broadly, five physicochemical properties of ENMs, seven abiotic factors related to aqueous conditions, six biotic properties related to biological organisms to support learning. Discontinuous 1,036 training examples in the dataset necessitated further split into four sub-subsets representing each of the responses selected or derived from the diverse endpoints investigated, reported, and collated from the scientific literature.

Chapter 5, 6 and 7 reports the modelling experiments and simulations performed using the sample dataset presented in Chapter 4, and interpretations of findings. Coding qualitative data and normalizing continuous data are important preparation procedures that determine the outcome from neural networks training. Application of a modified frequency-based coding scheme numerically coded nominal levels of categorical data to support training using the MLP algorithm that performs better with numerical data. The coding reduced the

dimensionality of potential categorical variables from 19, 36, and 42 to 2, 3, and 10, applied in Chapter 5, 6, and 7, respectively. Four trained models predicted the following,

1. Transfer of organic adsorptive on $nTiO_2$ surface and size evolution of the ENM in liquid-solid interactions in an aqueous environment. Two independent responses were the amount of organic matter adsorbed on the ENM and the hydrodynamic size of the ENM in the presence of organic matter.
2. Toxicity effects of the ENM to algae and water flea, *Daphnia magna* indicator organisms. Two independent relative effect responses were (i) inhibited biomass growth in algae, and (ii) inhibited reproduction in *D. magna.*

Chapters 5, 6 and 7 present the development of models, application of input perturbation and ranking methods to identify the relative importance of input variables and their contribution to predicting the responses. The models reported in Chapters 5 and 6 fitted to suitable input data forecast adsorbed mass and hydrodynamic size of ENMs that associated responses described in (1), and between responses in (1) and those in (2), all above. The significance of predicting behaviour kinetics' responses relates to the mobility, bioavailability, toxicity and management of ENMs in aquatic environment. The significance of the toxicity-based models underscores potential effects on population dynamics in the ecosystems.

8.2 Conclusions

There have been, and still are experimental, analytical, and modelling approaches pursued to investigate the behaviour kinetics and toxicity of ENMs an aqueous environment. Experimental studies have been concerned with generating knowledge on factors that would influence the fate of ENMs in aquatic environment and effective concentrations that induce toxicity effects in organisms for application in aquatic risk assessment. The analysis of environmental samples focuses on establishing the concentration of ENMs in different compartments, which are like large-scale models.

The three approaches mentioned above have separately anchored this study, which created connectivity between their disjointed aspects. One of the core themes was to consolidate available experimental data from multiple sources, simulate pseudo-continuous exposure systems and predict the responses in batch mode supervised learning using coexisting diverse factors. By doing so, a virtual continuum of three broad categories of factors governing behaviour kinetics and toxicity of ENMs in the aqueous environment affords the

understanding of complex interactions expected in the aquatic environment and contribution of each variable in explaining responses identified.

Another main theme was to apply proposed methods to approximate and aggregate several models using set criteria. The minimum validation set errors, testing set errors, variances of mean square errors, and R^2 evaluated the prediction performance of neural networks in learning from mixed data to extract vital information. The following is a list of notable outcomes from supervised training and simulations

- Despite some variations, approximated response outputs are in line with the current knowledge about the behaviour kinetics and toxicity of $nTiO_2$ in an aqueous environment.
- When using multiple inputs, the surface area concentration (SAC) is a primary variable in toxicity on *D. magna*, intermediate in adsorption process and secondary in aggregation and toxicity on algae owing to differences in size and concentration of $nTiO_2$ in datasets.
- The capping agents enhance size evolution of $nTiO_2$ and moderate toxicity effects on algae and *D. magna* depending on the effects of abiotic and biotic factors acting on physical properties of the ENM in an aqueous environment.
- The stability of protective capping agents will determine size evolution, potential mobility, bioavailability, and toxicity of ENMs in an aqueous environment.
- Data generated using surrogate organic and inorganic matter, and standard synthetic and purified water may not be sufficient for use in predicting behaviour kinetics and toxicity of ENMs in the aquatic environment.
- The concentration and speciation of mixed ions influence surface charge, behaviour and toxicity of ENMs different from mono ions. Besides, K^+ displayed pronounced effects on adsorption than Ca^{2+} that contrasts expected valence activity.
- The SAC does not represent size in principles governing aggregation kinetics of charged solids.
- Adsorption and aggregation processes likely to occur simultaneously in the natural environment do not follow a pattern and have contrasting influence on toxicity of $nTiO_2$.
- A nonlinear and non-specific relationship between adsorption, aggregation, and toxicity emphasizes complex interactions of ENMs with ligands, including organisms, determined by prevailing environmental conditions that require a good understanding of the systems for meaningful estimation.

- The model is limited to interpolated rather than extrapolated prediction of responses. However, generic input data evaluates MLP-NN model sensitivity, thus predictive power.

Coding and normalizing data in the dataset rather than in separate response data subsets ensured consistency in deploying and evaluating the sensitivity of the fitted models. Sensitivities presented in this study are static, with only one variable held constant or varied between the minimum and maximum range in the fitted models during experiments. The fitted models were not optimized because interpreting analyzed fixed quantitative scores of categorical variables representing unknown variables would be meaningless.

8.3 Contribution of the study

This study is relevant in environmental modelling. It is clear from Chapters 3 and 4 that application of MLP neural networks and secondary data in research are not new concepts. Nevertheless, contributions from this thesis include the following,

- New aspects in data compilation
    - Comprehensive data lists experimental designs, materials, methods, and findings.
    - Relevant data not reported in scientific publications derived using test protocols and other reference materials. Other data merged to generate new variables.
    - Relative toxicity responses derived as a percentage of reported effects created comparative increase and a decrease in toxicity effects having uniform units.
    - Similar properties of diverse aqueous media compiled for the first time to define components applicable in aquatic environment.
    - The quality of collated data determined based on criteria developed.
- Modelling using the MLP neural networks
    - The technique, applied for the first time in nanoecotoxicology, effectively predicts responses identified from endpoints of adsorption and aggregation kinetics, and toxicity effects of ENMs to indicator organisms.
    - Causal multiple physicochemical properties of ENMs, and environmental abiotic and biotic factors applied in the study simulate the aquatic environment and are difficult to measure in a single experimental study.
    - Training data used comprises new continuous and categorical input and response variables.

- The SAC integrates size, morphology, and concentration of ubiquitous ENMs and nanostructures and a potential predictor of their behaviour and toxicity.
- The performance of the fitted ensemble models in simulating responses using new data largely agrees with experimental findings whose data were not used in training.

- Applications
  - The MLP-NN model is explorative, predictive, and adaptable to external changes in input variables, thus worth considering as a tool in nanoecotoxicology research for risk management of ENMs.
  - The successful use of multiple interacting qualitative and quantitative factors shows potential use of the model in similar exposure systems to generalize phenomena.
  - Capable of extracting explanatory contribution of inputs to the prediction of responses in situations having associations not well understood.
  - The batch mode modelling approach demonstrates the model is usable in multiple stage dynamic systems, which can be coupled to develop mechanistic models that explain behaviour and toxicity of ENMs in an aquatic environment.

8.4 Recommendations and Future Work

The thesis proposed to collate data from accessible scientific reports and investigate the contribution of physicochemical properties of ENMs, and environmental abiotic and biotic factors to approximating behavior and toxicity responses in an aqueous environment. The research outputs targeted potential application of findings to an aquatic environment. As the research progressed, spaces for future work opened, which can strengthen data-driven forecasting. This Section proposes improvement and relevant areas for consideration in future work.

The first is the application of neural networks to predict behaviour kinetics and toxicity of notable ENMs that are potential pollutants in the aquatic environment, and simulate covariance effects on responses. Comparative studies will appraise the suitability of the model as a tool for environmental modelling to support research concerning ENMs pollutants in aquatic ecosystems. Second, considering numerous quantitative and qualitative variables, coding the latter type of data requires enhancement to reduce potential ordering or coupling with a different algorithm that handles categorical data in a better way.

Simultaneous measurement and reporting of behaviour kinetics of ENMs and toxicity endpoints may not be feasible. As a step towards reducing data ambiguities, the third proposal underscores comprehensive characterization of ENMs and properties of test water used during experiments. The descriptions of input variables would ultimately enhance consistency, interpretation, and application of findings. The final recommendation is the development and adoption of standard testing methods as well as consistent reporting of units of measurement for inputs and endpoints investigated to enable comparisons, conversions, and application in modelling.

Suggestions for future research work include the following,

- Investigate behaviour and toxicity of surface modified pristine ENMs used as ingredients in nanoproducts. Knowledge from the findings will be beneficial in understanding surface transformation, bioavailability, and toxicity effects of ENMs inferable to aquatic systems.

- Investigate partitioning constants of organic and inorganic ligands with ENMs in water usable as model predictors of ecological responses.

- Simulate proximate natural aquatic conditions at micro-, meso-, and macrocosm levels to minimize uncertainties and enhance the quality of data generated at bench scale. Models trained using data generated at bench-scale be validated using higher-level data to improve extrapolation of findings to the aquatic environment.

- Using environmentally reported concentrations of the ENMs, examine translational factors from minuscule-to-medium-to-pilot scale systems using dimensional analysis and similitude theories to support extrapolation of findings.

- Investigate effects of insufficient photon exposure to photosynthetic organisms, and photocatalytic-related toxicity, to distinguish effects of photocatalytic ENMs in the absence and presence of irradiation, respectively.

- Charged ions effectively represented micro- and macro elements in generic and natural water as a qualitative predictor in the model. Thus, the quantitative form of the variable requires investigation to determine explanatory contribution of individual elements to prediction of ENMs' behaviour and toxicity responses of ENMs.

References

Abbott, L.C. and Maynard, A.D. (2010). Exposure assessment approaches for engineered nanomaterials. *Risk Analysis*. 30 (11): 1634-1644.

Abbott, W.S. (1925). A method of computing the effectiveness of an insecticide. *Economic Entomology*. 18: 265-267.

Abraham, M.H. (1993). Scales of solute hydrogen-bonding: Their construction and application to physicochemical and biochemical processes. *Chemical Society Review*. 22 (5): 73-83.

Ackoff, R.L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*. 16 (1): 3-9.

Aguilera, K., Bishof, U., Karsten, D. and Hanelt, W.C. (2002). Seasonal variation in ecophysiological patterns in macroalgae from an Arctic fjord. II. Pigment accumulation and biochemical defence systems against high light stress. *Marine Biology*. 140 (6): 1087-1095.

Aiken, G.R., Hsu-Kim, H. and Ryan, J.N. (2011). Influence of dissolved organic matter on the environmental fate of metals, nanoparticles, and colloids. *Environmental Science and Technology*. 45 (8): 3196-3201.

Akin, C., Yi, J., Feldman, L.C., Durand, C., Hus, S.M., Li, A., Filler, M.A. and Shan, J.W. (2015). Contactless determination of electrical conductivity of one-dimensional nanomaterials by solution-based electro-orientation Spectroscopy. *ACS Nano*. 9 (5): 5405-5412.

Aldrich, C. (2002). *Exploratory analysis of metallurgical process data with neural networks and related methods*. Stellenbosch, CPT South Africa: Elsevier.

Alvarez, P.J.J., Colvin, V., Lead, J. and Stone, V. (2009). Research priorities to advance eco-responsible nanotechnology. *ACS Nano*. 3 (7): 1616-1619.

Anderson, W.C. (2005). Turbidity. In: *Techniques of Water Resources Investigations (TWRI) Book 9*. United States Geological Survey, pp. 1-55.

Antunes, S.C., Castro, B.B. and Goncalves, F. (2004). Effect of food level on the acute and chronic responses of daphnids to lindane. *Environmental Pollution*. 127 (3): 367-375.

Apul, O.G., Wang, Q., Shao, T. and Rieck, J.R. (2013). Predictive model development for adsorption of aromatic contaminants by multi-walled carbon nanotubes. *Environmental Science and Technology*. 47 (5): 2295-2303.

Araujo, P.Z., Morando, P.J. and Blesa, M.A. (2005). Interaction of catechol and gallic acid with titanium dioxide in aqueous suspensions. 1. Equilibrium studies. *Langmuir*. 21 (8): 3470-3474.

Arnqvist, G. and Wooster, D. (1995). Meta-analysis: synthesizing research findings in ecology and evolution. *Trends in Ecology and Evolution*. 10 (6): 236-240.

Aruoja, V., Dubourguier, H., Kasemets, K. and Kahru, A. (2009). Toxicity of nanoparticles of CuO, ZnO and TiO_2 to microalgae *Pseudokirchneriella subcapitata. Science of the Total Environment*. 407 (4): 1461-1468.

Aschberger, K., Gottardo, S., Amenta, V., Arena, M., Moniz, F.B., Bouwmeester, H., Brandhoff, P., Mech, A., Pesudo, L.Q., Rauscher, H., Schoonjans, R., Vettori, M.V. and Peters, R. (2014). Nanomaterials in Food - Current and Future Applications. In: *4th International conference on safe production and use of nanomaterials (Nanosafe2014)*. 2014, Physics: Conference Series 617, pp. 1-8.

Auffan, M., Pedeutour, M., Jérôme, R., Masion, A., Ziarelli, F., Borschneck, D., Chaneac, C., Botta, C., Chaurand, P., Labille, J. and Bottero, J.-Y. (2010). Structural degradation at the surface of a TiO_2-based nanomaterial used in cosmetics. *Environmental Science and Technology*. 44 (7): 2689-2694.

Baalousha, M. (2009). Aggregation and disaggregation of iron oxide nanoparticles: Influence of particle concentration, pH and natural organic matter. *Science of the Total Environment*. 407 (6): 2093-2101.

Babbie, E. and Mouton, J. (1998). *The practice of social research*. Cape Town, South Africa: Oxford University Press.

Banerjee, S., Gopal, J., Muraleedharan, P., Tyagi, A.. and Raj, B. (2006). Physics and chemistry of photocatalytic titanium dioxide: Visualization of bactericidal activity using atomic force microscopy. *Current Science*. 90 (10): 1378-1383.

Barnard, A.S. (2009). Modelling the relative stability of carbon nanotubes exposed to environmental adsorbates and air. *Physics: Condensed Matter - An Institute of Physics Journal*. 21 (14). p.p. 144205.

Barreto, W.J., Ando, R. A., Estevão, B.M. and Zanoni, K.P.D.S. (2012). Adsorption of caffeic acid on titanium dioxide: A spectroscopic study. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*. 92: 16-20.

Basheer, I. and Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing,

design, and application. *Microbiological Methods*. 43: 3-31.

Batley, G., Kirby, J.K. and Mclaughlin, M.J. (2013). Fate and risks of nanomaterials in aquatic and terrestrial environment. *Accounts of Chemical Research*. 46 (3): 854-862.

Battin, T.J., Kammer, F.V.D., Weilhartner, A., Ottofuelling, S. and Hoffmann, T. (2009). Nanostructured TiO_2: transport behavior and effects on aquatic microbial communities under environmental conditions. *Environmental Science and Technology*. 43 (21): 8098-8104.

Battiti, R. (1992). First- and second-order methods for learning: Between steepest descent and Newton's method. *Neural Computation*. 4 (2): 141-166.

Baughman, G.L. and Lassiter, R.R. (1978). Prediction of environmental pollutant concentration. In: J. Cairns, K. L. Dickson, and A. W. Maki (eds.). *Estimating the Hazard of Chemical Substances to Aquatic Life*. Philadelphia: American Society for Testing and Materials, pp. 35-54.

Baughman, R.H., Zakhidov, A.A. and de Heer, W.A. (2002). Carbon nanotubes-the route toward applications. *Science*. 297 (5582): 787-792.

Bayram, A., Kankal, M. and Onsoy, H. (2012). Estimation of suspended sediment concentration from turbidity measurements using artificial neural networks. *Environmental Monitoring and Assessment*. 184 (7): 4355-4365.

Beckett, R., Jue, Z. and Giddings, C.J. (1987). Determination of molecular weight distributions of fulvic and humic acids using flow field-flow fractionation. *Environmental Science and Technology*. 21 (3): 289-295.

Berger, T., Sterrer, M., Diwald, O. and Knözinger, E. (2005). Charge trapping and photoadsorption of O_2 on dehydroxylated TiO_2 nanocrystals-an electron paramagnetic resonance study. *ChemPhysChem*. 6 (10): 2104-2112.

Bhattacharjee, S., Elimelech, M. and Borkovec, M. (1998). DLVO Interaction between Colloidal Particles : Beyond Derjaguin's Approximation. *Croatica Chemica ACTA*. 71 (4): 5-8.

Bianco, A., Kostarelos, K. and Prato, M. (2005). Applications of carbon nanotubes in drug delivery. *Current Opinion in Chemical Biology*. 9 (6): 674-679.

Binnig, G. and Rohrer, H. (1987). Scanning tunneling microscopy- from birth to adolescence. *Reviews of Modern Physics*. 59 (3): 615-625.

Bishop, C.M. (1995). *Neural networks for pattern recognition*. 1st. Ed. Oxford, UK: Clanderon Press.

Bishop, C.M. (2006). *Pattern recognition and machine learning*. 2nd Ed. M. Jordan, J. Kleinberg, and B. Scholkopf (eds.). Cambridge, UK: Springer.

Blaser, S.A., Scheringer, M., Macleod, M. and Hungerbühler, K. (2007). Estimation of cumulative aquatic exposure and risk due to silver: Contribution of nano-functionalized plastics and textiles. *Science of the Total Environment*. 390 (2-3): 396-409.

Borm, P.J.A.A., Robbins, D., Haubold, S., Kuhlbusch, T., Fissan, H., Donaldson, K., Schins, R., Stone, V., Kreyling, W., Lademann, J., Krutmann, J., Warheit, D. and Oberdorster, E. (2006). The potential risks of nanomaterials: a review carried out for ECETOC. *Particle and Fibre Toxicology*. 35: 1-35.

Botta, C., Labille, J., Auffan, M., Borschneck, D., Miche, H., Cabié, M., Masion, A., Rose, J. and Bottero, J.-Y. (2011). TiO_2-based nanoparticles released in water from commercialized sunscreens in a life-cycle perspective: structures and quantities. *Environmental Pollution*. 159 (6): 1543-1550.

Bowman, D.M. and Hodge, G.A. (2007). A small matter of regulation: an international review of nanotechnology regulation. *The Colombia Science and Technology Law Review*. VIII: 1-36.

Boxall, A.B.A., Chaudhry, Q., Sinclair, C., Jones, A., Aitken, R., Jefferson, B. and Watts, C. (2007). *Current and future predicted environmental exposure to engineered nanoparticles*. York.

Boxall, A.B.A., Tiede, K. and Chaudhry, Q. (2007). Engineered nanomaterials in soils and water: How do they behave and could they pose a risk to human health? *Nanomedicine*. 2 (6): 919-927.

Brar, S.K., Verma, M., Tyagi, R.D. and Surampalli, R.Y. (2010). Engineered nanoparticles in wastewater and wastewater sludge-evidence and impacts. *Waste Management*. 30 (3): 504-520.

Breiman, L. (1996). Heuristics of instability in model selection. *The Annals of Statistics*. 24 (6): 2350-2383.

Breiman, L. (1996). Bagging Predictors. *Machine Learning*. 140: 123-140.

Broomhead, D.H. and Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems*. 2: 321-355.

Brown, G. (2004). *PhD Thesis: Diversity in neural network ensembles*. The University of Birmingham, United Kingdom.

Broyden, C.. G.. (1970). The convergence of a class of double rank minimization algorithms: 2. The new algorithm. *Journal of the Institute of Mathematics and Its Applications*. 6: 76-231.

Brunauer, S., Emmett, P.H. and Teller, E. (1938). Adsorption of gases in multimolecular layers. *American Chemical Society*. 60 (2): 309-319.

Buffle, J. and Leppard, G.G. (1995). Characterization of aquatic colloids and macromolecules. 1. Structure and behavior of colloidal material. *Environmental Science and Technology*. 29 (9): 2169-2175.

Buffle, J., Wilkinson, K.J. and Stoll, S. (1998). A Generalized Description of Aquatic Colloidal Interactions : The Three-colloidal Component Approach. *Environmental Science and Technology*. 32 (19): 2887-2899.

Burney, S.M.A., Jilani, T.A. and Ardil, C. (2007). A comparison of first and second order training algorithms for artificial neural networks. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*. 1 (1): 145-151.

Cadek, M., Coleman, J.N., Barron, V., Hedicke, K. and Blau, W.J. (2002). Morphological and mechanical properties of carbon-nanotube-reinforced semicrystalline and amorphous polymer composites. *Applied Physics Letters*. 81 (27): 5123-5125.

Cairns, J.Jr., Heath, A.G. and Parker, B.C. (1975). The effects of temperature upon the toxicity of chemicals to aquatic organisms. *Hydrobiologia*. 47 (1): 135-171.

Campos, B., Rivetti, C., Rosenkranz, P., María, J. and Barata, C. (2013). Effects of nanoparticles of TiO_2 on food depletion and life-history responses of *Daphnia magna. Aquatic Toxicology*. 130-131: 174-183.

Cao, G. (2004). *Nanostructures and nanomaterials: synthesis, properties and applications*. 1st Ed. J. Brinker (ed.). London: Imperial College Press.

Cardinale, B.J., Bier, R. and Kwan, C. (2012). Effects of TiO_2 nanoparticles on the growth and metabolism of three species of freshwater algae. *Nanoparticles Research*. 14 (913): 1-8.

Casals, E., Vazquez-Campos, S., Bastus, N.G. and Puntes, V. (2008). Distribution and potential toxicity of engineered inorganic nanoparticles and carbon nanostructures in biological systems. *Trends in Analytical Chemistry*. 27 (8).

Cassee, F.R., Van Balen, E.C., Singh, C., Green, D., Muijser, H., Weinstein, J. and Dreher, K. (2011). Exposure, health and ecological effects review of engineered nanoscale cerium and cerium oxide associated with its use as a fuel additive. *Critical Reviews in Toxicology*. 41 (3): 213-229.

Chekli, L., Roy, M., Tijing, L.D., Donner, E., Lombi, E. and Shon, H.K. (2015). Agglomeration behaviour of titanium dioxide nanoparticles in river waters: a multi-method approach combining light scattering and field-flow fractionation techniques. *Environmental Management*. 159: 135-142.

Chekli, L., Zhao, Y.X., Tijing, L.D., Phuntsho, S., Donner, E., Lombi, E., Gao, B.Y. and Shon, H.K. (2015). Aggregation behaviour of engineered nanoparticles in natural waters: Characterising aggregate structure using on-line laser light scattering. *Hazardous Materials*. 284: 190-200.

Chen, G., Liu, X. and Su, C. (2012). Distinct effects of humic acid on transport and retention of TiO_2 rutile nanoparticles in saturated sand columns. *Environmental Science and Technology*. 46 (13): 7142-7150.

Chen, K.L. and Elimelech, M. (2006). Aggregation and deposition kinetics of Fullerene (C_{60}) nanoparticles. *Langmuir*. 117 (18): 10994-11001.

Chen, K.L. and Elimelech, M. (2007). Influence of humic acid on the aggregation kinetics of fullerene (C60) nanoparticles in monovalent and divalent electrolyte solutions. *Colloid and Interface Science*. 309 (1): 126-134.

Chen, K.L. and Elimelech, M. (2008). Interaction of fullerene (C_{60}) nanoparticles with humic acid and alginate coated silica surfaces: Measurements, mechanisms, and environmental implications. *Environmental Science and Technology*. 42 (20). p.p. 7607-7614.

Chen, K.L., Mylon, S.E. and Elimelech, M. (2006). Aggregation kinetics of nanoparticles in monovalent and divalent electrolytes. *Environmental Science and Technology*. 40 (5): 1516-1523.

Chen, X. and Mao, S.S. (2007). Titanium dioxide nanomaterials: Synthesis, properties, modifications, and applications. *Chemical Reviews*. 107: 2891-2959.

Chen, X. and Schluesener, H.J. (2008). Nanosilver: a nanoproduct in medical application. *Toxicology letters*. 176 (1): 1-12.

Chen, Y., Huang, Y. and Li, K. (2012). Temperature effect on the aggregation kinetics of

CeO_2 nanoparticles in monovalent and divalent electrolytes. *Environmental and Analytical Toxicology*. 2 (7): 1-5.

Chin, Y., Alken, G. and Loughlins, E.O. (1994). Molecular weight, polydispersity, and spectroscopic properties of aquatic humic substances. *Environmental Science and Technology*. 28 (11): 1853-1858.

Chiou, C., Peters, L. and Freed, V. (1979). A physical concept of soil-water equilibria for nonionic organic compounds. *Science*. 206 (4420): 831-832.

Chowdhury, I., Cwiertny, D.M. and Walker, S. (2012). Combined factors influencing the aggregation and deposition of nano-TiO 2 in the presence of humic acid and bacteria. *Environmental Science and Technology*. 46 (13): 6968-6976.

Chowdhury, I., Hong, Y., Honda, R.J. and Walker, S.L. (2011). Mechanisms of TiO_2 nanoparticle transport in porous media: Role of solution chemistry, nanoparticle concentration, and flow rate. *Colloid and Interface Science*. 360 (2): 548-555.

Chowdhury, I., Walker, S.L. and Mylon, S.E. (2013). Aggregate morphology of nano-TiO_2: role of primary particle size, solution chemistry, and organic matter. *Environmental Science: Processes and Impacts*. 15 (1): 275-282.

Christian, P., Von der Kammer, F., Baalousha, M. and Hofmann, T. (2008). Nanoparticles: structure, properties, preparation and behaviour in environmental media. *Ecotoxicology*. 17: 326-343.

Church, R.M. (2001). The effective use of secondary data. *Learning and Motivation*. 33 (1): 32-45.

Clément, L., Hurel, C. and Marmier, N. (2013). Toxicity of TiO_2 nanoparticles to cladocerans, algae, rotifers, and plants - Effects of size and crystalline structure. *Chemosphere*. 90 (3): 1083-1090.

Cohen, J., Cohen, P., West, S.G. and Aiken, L.S. (2003). *Applied multiple regression / correlation analysis for the behavioral sciences*. 3rd Ed. Lawrence Erlbaum Associates.

Coleman, J.N., Khan, U., Blau, W.J. and Gun'ko, Y.K. (2006). Small but strong: A review of the mechanical properties of carbon nanotube-polymer composites. *Carbon*. 44 (9): 1624-1652.

Crane, M., Handy, R.D., Garrod, J.F. and Owen, R. (2008). Ecotoxicity test methods and environmental hazard assessment for engineered nanoparticles. *Ecotoxicology*. 17 (5): 421-37.

Cupi, D., Hartmann, N.B. and Baun, A. (2015). The influence of natural organic matter and aging on suspension stability in guideline toxicity testing of silver, zinc oxide, and titanium dioxide nanoparticles with *Daphnia magna*. *Environmental Toxicology and Chemistry*. 34 (3): 497-506.

Cybenko, G. (1989). Degree of approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*. 2 (4): 303-314.

Dąbrowski, A. (2001). Adsorption-from theory to practice. *Advances in Colloid and Interface Science*. 93 (1-3): 135-224.

Dabrunz, A., Duester, L., Prasse, C., Seitz, F., Rosenfeldt, R., Schilde, C., Schaumann, G.E. and Schulz, R. (2011). Biological surface coating and molting inhibition as mechanisms of TiO_2 nanoparticle toxicity in *Daphnia magna*. *PLoS One*. 6 (5): 1-7.

Dakota Matrix Minerals (2012). *Crystallography and minerals arranged by crystal form*. http://www.dakotamatrix.com/. [Accessed: 12 July 2012].

Danish Consumer Council (2015). *The Nanodatabase*. [Online]. 2015. Available from: http://nanodb.dk/. [Accessed: 6 November 2015].

Das, P., Xenopoulos, M.A. and Metcalfe, C.D. (2013). Toxicity of silver and titanium dioxide nanoparticle suspensions to the aquatic invertebrate, *Daphnia magna*. *Bulletin of Environmental Contamination and Toxicology*. 91 (1): 76-82.

Davidovits, J. (1991). Geopolymers: Inorganic polymeric new materials. *Thermal Analysis*. 37 (8): 1633-1656.

Davis, J.A. (1982). Adsorption of natural dissolved organic matter at the oxide/water interface. *Geochimica et Cosmochimica Acta*. 46: 2381-2393.

Davis, M.J. (2010). Contrast coding in multiple regression analysis: strengths, weaknesses, and utility of popular coding structures. *Data Science*. 8: 61-73.

Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap methods and their application*. 1st Ed. Cambridge University Press.

Dedecker, A.P., Goethals, P.L., D'heygere, T., Gevrey, M., Lek, S. and De Pauw, N. (2005). Application of artificial neural network models to analyse the relationships between gammarus pulex L. (Crustacea, Amphipoda) and river characteristics. *Environmental Monitoring and Assessment*. 111 (1-3): 223-241.

Derjaguin, B. and Landau, L. (1941). Theory of the stability of strongly charged lyophobic

sols and of the adhesion of strongly charged particles in solution of electrolytes. *Acta Physicochim USSR*. 14: 633-662.

Dhawan, A. and Sharma, V. (2010). Toxicity assessment of nanomaterials : methods and challenges. *Analytical and Bioanalytical Chemistry*. 398: 589-605.

Diebold, U. (2003). The surface science of titanium dioxide. *Surface Science Reports*. 48: 53-229.

Dimopoulos, I., Chronopoulos, J., Chronopoulou-Sereli, A. and Lek, S. (1999). Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in Athens city (Greece). *Ecological Modelling*. 120 (2-3): 157-165

Domingos, E.F., Tufenkji, N. and Wilkinson, K.J. (2009). Aggregation of titanium dioxide nanoparticles: role of a fulvic acid. *Environmental Science and Technology*. 43 (5): 1282-1286.

Douglas, R.W. and Rippey, B. (2000). The random redistribution of sediment by wind in a lake. *Limnology and Oceanography*. 45 (3): 686-694.

Dumont, E., Johnson, A.C., Keller, V.D.J. and Williams, R.J. (2015). Nano silver and nano zinc-oxide in surface waters-Exposure estimation for Europe at high spatial and temporal resolution. *Environmental Pollution*. 196: 341-349.

Duncan, T. V. (2011). Applications of nanotechnology in food packaging and food safety: barrier materials, antimicrobials and sensors. *Colloid and Interface Science*. 363 (1): 1-24.

Ebert, D. (2005). *Ecology, Epidemiology, and Evolution of Parasitism in Daphnia (Internet)*. 1st Ed. [Online]. Bethesda (MD): National Library of Medicine (US): National Center for Biotechnology Information (NCBI). Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Books.

Eby, G.N. (2004). Carbon chemistry. In: *Principles of environmental geochemistry*. Thomson-Brooks / Cole-Thomson Learning.

ECB (2003). *Technical guidance document on risk assessment Part II*.

ECHA. *Appendix R8-15 Recommendations for nanomaterials applicable to Chapter R.8 Characterisation of dose [concentration] - response for human health*. (Appendix R8-15).

ECHA (2015). *Regulations: nanomaterials*. [Online]. 2015. Available from:

http://echa.europa.eu/regulations/nanomaterials. [Accessed: 7 November 2015].

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*. 7 (1): 1-26.

El Badawy, A.M., Silva, R.G., Morris, B., Scheckel, K.G., Suidan, M.T. and Tolaymat, T.M. (2011). Surface charge-dependent toxicity of silver nanoparticles. *Environmental Science and Technology*. 45 (1): 283-287.

Elendt, B.P. and Bias, W.R. (1990). Trace nutrient deficiency in *Daphnia magna* cultured in standard medium for toxicity testing. Effects of the optimization of culture conditions on life history parameters of *D. magna. Water Research*. 24 (9): 1157-1167.

Elimelech, M., Gregory, J., Jia, X. and Williams, R.A. (1995). *Particle Deposition and Aggregation: Measurement, Modelling and Simulation*. 2nd Ed. R. A. Williams (ed.). Oxford: Butterworth-Heinemann Ltd.

EPA (2015). *Nanoscale materials; chemical substances when manufactured, imported, or processed as nanoscale materials; reporting and recordkeeping requirements*. [Online]. 2015. Available from: http://yosemite.epa.gov/opei/rulegate.nsf/byRIN/2070-AJ54. [Accessed: 7 November 2015].

Erhayem, M. and Sohn, M. (2014a). Effect of humic acid source on humic acid adsorption onto titanium dioxide nanoparticles. *Science of the Total Environment*. 470-471: 92-98.

Erhayem, M. and Sohn, M. (2014b). Stability studies for titanium dioxide nanoparticles upon adsorption of Suwannee River humic and fulvic acids and natural organic matter. *Science of the Total Environment*. 468-469: 249-257.

EU (2011). *Commission recommendation on the definition of nanomaterial*. (2011/696/EU).

Fabian, E., Landsiedel, R., Ma-Hock, L., Wiench, K., Wohlleben, W. and Ravenzwaay, B. van (2008). Tissue distribution and toxicity of intravenously administered titanium dioxide nanoparticles in rats. *Archives of Toxicology*. 82 (3): 151-157.

Falkenhainer, B.C. and Michalski, R.S. (1986). Integrating quantitative and qualitative discovery: The ABACUS system. *Machine Learning*. 1 (4): 367-401.

Famili, A., Shen, W., Weber, R. and Simoudis, E. (1997). Data Preprocessing and Intelligent Data Analysis. *Intelligent Data Analysis*. 1 (1-4): 3-23.

Fan, Z. and Lu, J.G. (2005). Zinc oxide nanostructures: synthesis and properties. *Nanoscience and Nanotechnology*. 5 (10): 1561-1573.

Faure, B., Salazar-Alvarez, G., Ahniyaz, A., Villaluenga, I., Berriozabal, G., De Miguel, Y.R. and Bergström, L. (2013). Dispersion and surface functionalization of oxide nanoparticles for transparent photocatalytic and UV-protecting coatings and sunscreens. *Science and Technology of Advanced Materials*. 14 (2): 023001-023024.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*. 17 (3): 37-54.

Feynman, R.P. (1992). There's plenty of room at the bottom. *Microelectromechanical Systems*. 1 (1): 60-66.

Finnegan, M.P., Zhang, H. and Banfield, J.F. (2007). Phase stability and transformation in titania nanoparticles in aqueous solutions dominated by surface energy. *The Physical Chemistry C*. 111 (5): 1962-1968.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics. 7 (2): 179-188.

Fissan, H., Neumann, S., Trampe, A., Pui, D.V.. and Shin, W.. (2006). Rationale and principle of an instrument measuring lung deposited nanoparticle surface area. *Nanotechnology and Occupational Health*. 9: 53-59.

Fitkov-Norris, E., Vahid, S. and Hand, C. (2012). Evaluating the impact of categorical data encoding and scaling on neural network classification performance: The case of repeat consumption of identical cultural goods. In: C. Jayne, S. Yue, and L. Iliadis (eds.). *Engineering Applications of Neural Networks: 13th International Conference, EANN 2012, London, UK, September 2012, Proceedings*. Springer-Verlag Berlin Heidelberg, pp. 343-352.

Fletcher, N. and Bartholomaeus, A. (2011). Regulation of nanotechnologies in food in Australia and New Zealand. *International Food Risk Analysis Journal*. 2 (2): 33-40.

Fletcher, R. (1970). A new aproach to variable metric algorithm. *The Computer Journal*. 13 (3): 317-322.

Fletcher, R. and Reeves, C.M. (1964). Function minimization by conjugate gradients. *The Computer Journal*. 7 (2) ppp. 149-154.

Forough, M. and Farhadi, K. (2010). Biological and green synthesis of silver nanoparticles. *Engineering and Environmental Science*. 34: 281-287.

Fortner, J.D., Lyon, D.Y., Sayes, C.M., Boyd, A.M., Falkner, J.C., Hotze, E.M., Alemany, L.B., Tao, Y.J., Guo, W., Ausman, K.D., Colvin, V.L. and Hughes, J.B. (2005). C60 in

water: Nanocrystal formation and microbial response. *Environmental Science and Technology*. 39 (11): 4307-4316.

Francis, L. (2001). *Neural Networks Demystified*. p.p. 289.

French, R.A., Jacobson, A.R., Kim, B., Isley, S.L., Penn, R.L. and Baveye, P.C. (2009). Influence of ionic strength, pH, and cation valence on aggregation kinetics of titanium dioxide nanoparticles. *Environmental Science and Technology*. 43 (5): 1354-1359.

Freund, Y. and Schapire, R.E. (1997). A desicion-theoretic generalization of on-line learning and an application to boosting. *Computer and System Sciences*. 55 (1): 119-139.

Friedman, J.H. (1997). On bias, variance, 0/1-loss, and the curse-of-dimensionality. *Data Mining and Knowledge Discovery*. 77 (1): 55-77.

Fu, L., Hamzeh, M., Dodard, S., Zhao, Y.H. and Sunahara, G.I. (2015). Effects of TiO_2 nanoparticles on ROS production and growth inhibition using freshwater green algae pre-exposed to UV irradiation. *Environmental Toxicology and Pharmacology*. 39 (3): 1074-1080.

Fujishima, A., Rao, T.N. and Tryk, D.A. (2000). Titanium dioxide photocatalysis. *Photochemistry and Photobiology C: Photochemistry Reviews*. 1 (1): 1-21.

Gensemer, R.W. and Playle, R.C. (1999). The bioavailability and toxicity of aluminum in aquatic environment. *Critical Reviews in Environmental Science and Technology*. 29 (4): 315-450.

Gevrey, M., Dimopoulos, I. and Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*. 160: 249-264.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*. 84 (3): 279-325.

Giesche, H. (2006). Mercury porosimetry: a general (practical) overview. *Particle and Particle Systems Characterization*. 23 (1): 1-11.

Goh, A.T.C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*. 9 (3): 143-151.

Goldberg, E., Scheringer, M., Bucheli, T.D. and Hungerbühler, K. (2015). Prediction of nanoparticle transport behavior from physicochemical properties: machine learning transport models. *Environmental Science: Nano*. 2: 352-360.

Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. *Mathematics of Computation*. 24 (109): 23-26.

Goldman, J.C., Middlebrooks, J.E. and Toerien, D.F. (1971). *The effect of carbon on algal growth-its relationship to eutrophication*. Utah, USA.

Goodwin, J. (2004). *Colloids and interfaces with surfactants and polymers - an introduction*. 1st Ed. John Wiley and Sons Inc.

Gottschalk, F., Kost, E. and Nowack, B. (2013). Engineered nanomaterials in waters and soils: a risk quantification based on probabilistic exposure and effect modeling. *Environmental Toxicology and Chemistry*. 32 (6): 1-29.

Gottschalk, F., Lassen, C., Kjoelholt, J. and Christensen, F. (2015). Modeling flows and concentrations of nine engineered nanomaterials in the Danish environment. *International Environmental Reseach and Public Health*. 12: 5581-5602.

Gottschalk, F., Ort, C., Scholz, R.W. and Nowack, B. (2011). Engineered nanomaterials in rivers-Exposure scenarios for Switzerland at high spatial and temporal resolution. *Environmental Pollution*. 159 (12): 3439-3445.

Gottschalk, F., Scholz, R.W. and Nowack, B. (2010). Probabilistic material flow modeling for assessing the environmental exposure to compounds: Methodology and an application to engineered nano-TiO_2 particles. *Environmental Modelling and Software*. 25 (3): 320-332.

Gottschalk, F., Sonderer, T., Scholz, R.W. and Nowack, B. (2009). Modeled environmental concentrations of engineered (TiO_2, Ag, CNT, ZnO, Fullerenes) for different regions. *Environmental Science and Technology*. 43 (24): 9216-9222.

Gottschalk, F., Sun, T. and Nowack, B. (2013). Environmental concentrations of engineered nanomaterials: Review of modeling and analytical studies. *Environmental Pollution*. 181: 287-300.

Grassian, V.H. (2008). When size really matters: size-dependent properties and surface chemistry of metal and metal oxide nanoparticles in gas and liquid phase environment. *Physical Chemistry*. 112: 18303-18313.

Griessbach, E.F.C. and Lehmann, R.G. (1999). Degradation of polydimethylsiloxane fluids. *Chemosphere*. 38 (6): 1461-1468.

Griffitt, R.J., Luo, J., Gao, J., Bonzongo, J.-C. and Barber, D.S. (2008). Effects of particle composition and species on toxicity of metallic nanomaterials in aquatic organisms.

Environmental Toxicology and Chemistry. 27 (9): 1972-1978.

Gruère, G.P. (2011). Labeling nano-enabled consumer products. *Nano Today*. 6 (2): 117-121.

Gu, B., Schmitt, J., Chen, Z., Liang, L. and McCarthy, J.F. (1995). Adsorption and adsorption of different organic matter fractions on iron oxide. *Geochimica et Cosmochimica Acta*. 59 (2): 219-229.

Guerrero-Garca, G.I., González-Mozuelos, P. and Olvera De La Cruz, M. (2011). Potential of mean force between identical charged nanoparticles immersed in a size-asymmetric monovalent electrolyte. *Chemical Physics*. 135 (16): 1-10.

Hagan, M.T., Demuth, H.B., Beale, M.H. and De Jesús, O. (2014). *Neural Network Design*. Martin Hagan.

Hager, W.W. and Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. *Pacific Journal of Optimization*. 2 (1): 35-58.

Hamaker, J.W. and Thompson, J.M. (1972). Adsorption. In: C. A. I. Goring and J. W. Hamaker (eds.). *Organic chemicals in the soil environment*. New York, USA: Marcel Dekker, Inc., New York, pp. 49-143.

Han, J., Cai, Y. and Cercone, N. (1993). Data-driven discovery of quantitative rules in relational databases. *IEEE Transactions on Knowledge and Data Engineering*. 5 (1): 29-40.

Han, J., Kamber, M. and Pei, J. (2012). *Data Mining: Concepts and Techniques*. 3rd Ed. Waltham, MA, USA: Morgan Kaufmann.

Hanaor, D.A.H. and Sorrell, C.C. (2011). Review of the anatase to rutile phase transformation. *Material Science*. 46 (4): 855-874.

Handy, R.D., van den Brink, N., Chappell, M., Mühling, M., Behra, R., Dušinská, M., Simpson, P., Ahtiainen, J., Jha, A.N., Seiter, J., Bednar, A., Kennedy, A., Fernandes, T.F. and Riediker, M. (2012). Practical considerations for conducting ecotoxicity test methods with manufactured nanomaterials: what have we learnt so far? *Ecotoxicology*. 21 (4): 933-972.

Hansen, F.S., Larsen, B.H., Olsen, S.I. and Baun, A. (2007). Categorization framework to aid hazard identification of nanomaterials. *Nanotoxicology*. 1 (3): 243-250.

Hariharan, C. (2006). Photocatalytic degradation of organic contaminants in water by ZnO

nanoparticles: Revisited. *Applied Catalysis A: General*. 304: 55-61.

Hartmann, N.B., Von der Kammer, F., Hofmann, T., Baalousha, M., Ottofuelling, S. and Baun, A. (2010). Algal testing of titanium dioxide nanoparticles: testing considerations, inhibitory effects and modification of cadmium bioavailability. *Toxicology*. 269 (2-3): 190-197.

Hastie, T., Robert, T. and Jerome, F. (2008). *The elements of statistical learning: data mining, inference, and prediction*. Second. Stanford, Los Angeles, California: Springer.

Hauri, J.F. and Niece, B.K. (2011). Leaching of silver from silver-impregnated food storage containers. *Chemical Education*. 88 (10): 1407-1409.

Haykin, S. (1998). *Neural networks: a comprehensive foundation*. 2nd. Ed. Ontario, Canada: Pearson Education, Inc.

Hecht-Nielsen, R. (1990). *Neurocomputing*. Boston, MA, USA: Addison-Wesley.

Heinlaan, M., Ivask, A., Blinova, I., Dubourguier, H.-C. and Kahru, A. (2008). Toxicity of nanosized and bulk ZnO, CuO and TiO_2 to bacteria Vibrio fischeri and crustaceans *Daphnia magna* and *Thamnocephalus platyurus. Chemosphere*. 71 (7): 1308-1316.

Helland, A., Scheringer, M., Siegrist, M., Kastenholz, H.G., Wiek, A. and Scholz, R.W. (2008). Risk assessment of engineered nanomaterials: a survey of industrial approaches. *Environmental Science and Technology*. 42 (2): 640-646.

Hendren, C.O., Lowry, M., Grieger, K.D., Money, E.S., Johnston, J.M., Wiesner, M.R. and Beaulieu, S.M. (2013). Modeling approaches for characterizing and evaluating environmental exposure to engineered nanomaterials in support of risk-based decision making. *Environmental Science and Technology*. 47: 1190-1205.

Hendren, C.O., Mesnard, X., Dröge, J. and Wiesner, M.R. (2011). Estimating production data for five engineered nanomaterials as a basis for exposure assessment. *Environmental Science and Technology*. 45 (7): 2562-2569.

Hendren, O.C., Badireddy, A.R., Casman, E. and Wiesner, M.R. (2013). Modeling nanomaterial fate in wastewater treatment : Monte Carlo simulation of silver nanoparticles (nano-Ag). *Science of the Total Environment*. 449: 418-425.

Hestenes, M.R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Research of the National Bureau of Standards*. 49 (6): 409-436.

Hornik, K. (1989). Multilayer feedforward networks are universal approximators. *Neural*

Networks. 2 (5): 359-366.

Hornyak, G.L., Tibbals, H.F., Dutta, J. and Moore, J.J. (2009). *Introduction to nanoscience and nanotechnology*. 1st Ed. Boca Raton: CRC Press.

Horst, A.M., Neal, A.C., Mielke, R.E., Sislian, P.R., Suh, W.H., Stucky, G.D. and Holden, P.A. (2010). Dispersion of TiO_2 nanoparticle agglomerates by *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*. 76 (21): 7292-7298.

Hou, W.-C., Westerhoff, P. and Posner, J.D. (2013). Environmental Science : Processes and Impacts Biological accumulation of engineered nanomaterials : a. *Environmental Science: Processes and Impacts*. 15: 103-122.

Hristovski, K.D., Westerhoff, P.K. and Posner, J.D. (2011). Octanol-water distribution of engineered nanomaterials. *Environmental Science and Health, Part A*. 46 (6): 37-41.

Hug, S.J. and Bahnemann, D. (2006). Infrared spectra of oxalate, malonate and succinate adsorbed on the aqueous surface of rutile, anatase and lepidocrocite measured with in situ ATR-FTIR. *Journal of Electron Spectroscopy and Related Phenomena*. 150 (2-3): 208-219.

Hull, M.S., Kennedy, A.J., Steevens, J.A., Bednar, A.J., Weiss, C.A. and Vikesland, P.J. (2009). Release of metal impurities from carbon nanomaterials influences aquatic toxicity. *Environmental Science and Technology*. 43 (11): 4169-4174.

Hulscher, T.E. Ten and Cornelissen, G. (1996). Effect of temperature adsorption on adsorption equilibrium and kinetics of organic micropollutants - a review. *Chemosphere*. 32 (4): 609-626.

Hund-Rinke, K. and Simon, M. (2006). Ecotoxic effect of photocatalytic active nanoparticles (TiO_2) on algae and daphnids. *Environmental Science Pollution Research*. 13 (4): 225-232.

Hunter, R.J. (1981). *Zeta potential in colloid science: principles and applications*. London, UK: Academic Press.

Huynh, K.A. and Chen, K.L. (2011). Aggregation kinetics of citrate and polyvinylpyrrolidone coated silver nanoparticles in monovalent and divalent electrolyte solutions. *Environmental Science and Technology*. 45 (13): 5564-5571.

Iswarya, V., Bhuvaneshwari, M., Alex, S.A., Iyer, S., Chaudhuri, G., Chandrasekaran, P.T., Bhalerao, G.M., Chakravarty, S., Raichur, A.M., Chandrasekaran, N. and Mukherjee, A. (2015). Combined toxicity of two crystalline phases (anatase and rutile) of titania

nanoparticles towards freshwater microalgae: *Chlorella sp*. *Aquatic Toxicology*. 161: 154-169.

Jackson, P., Jacobsen, N.R., Baun, A., Birkedal, R., Kühnel, D., Jensen, K.A., Vogel, U. and Wallin, H. (2013). Bioaccumulation and ecotoxicity of carbon nanotubes. *Chemistry Central Journal*. 7 (154): 1-21.

Jacobasch, C., Völker, C., Giebner, S., Völker, J., Alsenz, H., Potouridis, T., Heidenreich, H., Kayser, G., Oehlmann, J. and Oetken, M. (2014). Long-term effects of nanoscaled titanium dioxide on the cladoceran *Daphnia magna* over six generations. *Environmental Pollution*. 186: 180-186.

Jacobs, J.F., van de Poel, I. and Osseweijer, P. (2010). Sunscreens with titanium dioxide ($TiO(_2)$) nano-particles: A societal experiment. *Nanoethics*. 4 (2): 103-113.

Jacobs, R.A. and Jordan, M.I. (1991). Adaptive mixtures of local experts. *Neural Computation*. 3: 79-87.

Jager, T. (2011). Some good reasons to ban ECx and related concepts in ecotoxicology. *Environmental Science and Technology*. 45 (19): 8180-8181.

Jaroniec, M., Lu, X. and Richard, M. (1991). Correlation between the fractal dimension and the microporous structure of a solid. *Chemical Monthly*. 122 (8-9): 577-584.

Jensen, H., Joensen, K.D., Jørgensen, J.-E., Pedersen, J.S. and Søgaard, G. (2004). Characterization of nanosized partly crystalline photocatalysts. *Nanoparticle Research*. 6 (5): 519-526.

Ji, J., Long, Z. and Lin, D. (2011). Toxicity of oxide nanoparticles to the green algae *Chlorella sp. Chemical Engineering Journal*. 170 (2-3): 525-530.

Jiang, J., Oberdörster, G. and Biswas, P. (2009). Characterization of size, surface charge, and agglomeration state of nanoparticle dispersions for toxicological studies. *Nanoparticle Research*. 11: 77-89.

Johnson, A., Cisowska, I., Jürgens, M., Keller, V., Lawlor, A. and Williams, R. (2011). *Exposure assessment for engineered silver nanoparticles throughout the rivers of England and Wales (CB0433)*.

Johnson, A.C., Bowes, M.J., Crossley, A., Jarvie, H.P., Jurkschat, K., Jürgens, M.D., Lawlor, A.J., Park, B., Rowland, P., Spurgeon, D., Svendsen, C., Thompson, I.P., Barnes, R.J., Williams, R.J. and Xu, N. (2011). An assessment of the fate, behaviour and environmental risk associated with sunscreen TiO_2 nanoparticles in UK field scenarios.

Science of the Total Environment. 409 (13): 2503-2510.

Johnson, A.C. and Park, B. (2012). Predicting contamination by the fuel additive cerium oxide engineered nanoparticles within the United Kingdom and the associated risks. *Environmental Toxicology and Chemistry*. 31 (11): 2582-2587.

Jonczyk, E. and Gilron, G. (2005). Acute and chronic toxicity testing with *Daphnia Sp*. In: C. Blaise and J.-F. Férard (eds.). *Small-scale freshwater toxicity investigations: Toxicity test methods*. Dordrecht, Netherlands: Springer, pp. 337-393.

Jovanović, B. (2014). Critical review of public health regulations of titanium dioxide, a human food additive. *Integrated Environmental Assessment and Management*. 9999 (9999): 1-11.

Kaegi, R., Sinnet, B., Zuleeg, S., Hagendorfer, H., Mueller, E., Vonbank, R., Boller, M. and Burkhardt, M. (2010). Release of silver nanoparticles from outdoor facades. *Environmental Pollution*. 158 (9): 2900-2905.

Kaegi, R., Ulrich, A., Sinnet, B., Vonbank, R., Wichser, A., Zuleeg, S., Simmler, H., Brunner, S., Vonmont, H., Burkhardt, M. and Boller, M. (2008). Synthetic TiO_2 nanoparticle emission from exterior facades into the aquatic environment. *Environmental Pollution*. 156 (2): 233-239.

Kaegi, R., Voegelin, A., Sinnet, B., Zuleeg, S., Hagendorfer, H., Burkhardt, M. and Siegrist, H. (2011). Behavior of metallic silver nanoparticles in a pilot wastewater treatment plant. *Environmental Science and Technology*. 45 (9): 3902-3908.

Kah, M., Zhang, X., Jonker, M.T.O. and Hofmann, T. (2011). Measuring and Modeling Adsorption of PAHs to Carbon Nanotubes Over a Six Order of Magnitude Wide Concentration Range. *Environmental Science and Technology*. 45: 6011-6017.

Kahru, A. and Dubourguier, H.-C. (2010). From ecotoxicology to nanoecotoxicology. *Toxicology*. 269 (2-3): 105-119.

Kamlet, M.J., Abboud, J.-L.M., Abraham, M.H. and Taft, R.W. (1983). Linear solvation energy relationships. 23. A comprehensive collection of the Solvatochromic parameters, dipolarity, alpha, and beta and some methods for simplifying the generalized Solvatochromic equation. *Organic Chemistry*. 48 (13): 2877-2887.

Von Der Kammer, F., Ferguson, L.P., Holden, P.A., Masion, A., Rogers, K.R., Klaine, S.J., Koelmans, A.A., Horne, N. and Unrine, J.M. (2012). Analysis of engineered nanomaterials in complex matrices (environment and biota): general considerations and

conceptual case studies. *Environmental Toxicology and Chemistry*. 31 (1): 32-49.

Von Der Kammer, F., Ottofuelling, S. and Hofmann, T. (2010). Assessment of the physico-chemical behavior of titanium dioxide nanoparticles in aquatic environment using multi-dimensional parameter testing. *Environmental Pollution*. 158 (12): 3472-3481.

Kandegedara, A. and Rorabacher, D.B. (1999). Noncomplexing tertiary amines as 'better' buffers covering the range of pH 3-11. Temperature dependence of their acid dissociation constants. *Analytical Chemistry*. 71 (15): 3140-3144.

Kaplan, B. and Dduchon, D. (1988). Combining qualitative and quantitative methods in information systems: A case study. *Management Information System Research Center*. 12 (4): 571-586.

Keller, A.A. and Lazareva, A. (2014). Predicted Releases of Engineered Nanomaterials: From Global to Regional to Local. *Environmental Science and Technology Letters*. 1 (1): 65-70.

Keller, A.A., Mcferran, S., Lazareva, A. and Suh, S. (2013). Global life cycle releases of engineered nanomaterials. *Nanoparticle Research*. 15 (1692-1694): 1-17.

Keller, A.A., Wang, H., Zhou, D., Lenihan, H.S., Cherr, G., Cardinale, B.J., Miller, R. and Ji, Z. (2010). Stability and aggregation of metal oxide nanoparticles in natural aqueous matrices. *Environmental Science and Technology*. 44 (6): 1962-1967.

Khan, M.A.Q. and Khan, M.A. (2008). *Effect of temperature on water flea Daphnia magna (Crustacea: Cladocera)*. [Online]. Available from: Nature Precedings <http://hdl.handle.net/10101/npre.2008.1909.1> (2008).

Khosravi, K., Hoque, M.E., Dimock, B., Hintelmann, H. and Metcalfe, C.D. (2012). A novel approach for determining total titanium from titanium dioxide nanoparticles suspended in water and biosolids by digestion with ammonium persulfate. *Analytica Chimica Acta*. 713: 86-91.

Kim, J.I., Park, H.-G., Chang, K.-H., Nam, D.H. and Yeo, M.-K. (2016). Trophic transfer of nano-TiO_2 in a paddy microcosm: A comparison of single-dose versus sequential multi-dose exposures. *Environmental Pollution*. 212: 316-324.

Kim, K.T., Klaine, S.J., Cho, J., Kim, S. and Kim, S.D. (2010). Oxidative stress responses of *Daphnia magna* exposed to TiO_2 nanoparticles according to size fraction. *Science of the Total Environment*. 408 (10): 2268-2272.

Kim, K.-T., Klaine, S.J. and Kim, S.D. (2014). Acute and chronic response of *Daphnia*

magna exposed to TiO_2 nanoparticles in agitation system. *Bulletin of Environmental Contamination and Toxicology*. (2006): 1-5.

Kiser, M.A., Westerhoff, P., Benn, T., Wang, Y., Pérez-Rivera, J. and Hristovski, K. (2009). Titanium nanomaterial removal and release from wastewater treatment plants. *Environmental Science and Technology*. 43 (17): 6757-6763.

Kissa, E. (1999). *Dispersions: Characterization, testing, and measurement*. 1st Ed. New York, USA: CRC Press.

Klaine, S.J., Alvarez, P.J.J., Batley, Graeme, E., Fernandes, T.F., Handy, R.D., Lyon, D.Y., Mahendra, S., McLaughlin, M.J. and Lead, J.R. (2008). Nanomaterials in the environment: Behaviour, fate, bioavailability, and effects. *Environmental Toxicology and Chemistry*. 27 (9): 1825-1851.

Koelmans, A.A., Nowack, B. and Wiesner, M.R. (2009). Comparison of manufactured and black carbon nanoparticle concentrations in aquatic sediments. *Environmental Pollution*. 157 (4): 1110-1116.

Köhler, A.R., Soma, C., Hellanda, A. and Gottschalk, F. (2008). Studying the potential release of carbon nanotubes throughout the application life cycle. *Cleaner Production*. 16 (8-9): 927-937.

Kulacki, K.J. and Cardinale, B.J. (2012). Effects of nano-titanium dioxide on freshwater algal population dynamics. *PloS One*. 7 (10). p.p. e47130 (pp 7).

Kunicki-Goldfinger, J.J., Freestone, I.C., McDonald, I., Hobot, J.A., Gilderdale-Scott, H. and Ayers, T. (2014). Technology, production and chronology of red window glass in the medieval period - rediscovery of a lost technology. *Archaeological Science*. 41: 89-105.

Labille, J., Feng, J., Botta, C., Borschneck, D., Sammut, M., Cabie, M., Auffan, M., Rose, J. and Bottero, J.-Y. (2010). Aging of TiO_2 nanocomposites used in sunscreen. Dispersion and fate of the degradation products in aqueous environment. *Environmental Pollution*. 158 (12): 3482-3489.

Lan, Q., Liu, H., Li, F. bai, Zeng, F. and Liu, C. shuai (2011). Effect of pH on pentachlorophenol degradation in irradiated iron/oxalate systems. *Chemical Engineering Journal*. 168 (3): 1209-1216.

Larson, I. and Attard, P. (2000). Surface charge of silver iodide and several metal oxides. Are all surfaces Nernstian? *Colloid and Interface Science*. 227: 152-163.

Lebrette, S., Pagnoux, C. and Abélard, P. (2004). Stability of aqueous TiO_2 suspensions:

influence of ethanol. *Colloid Interface Science*. 280 (2): 400-408.

Leenheer, J.A., Rostad, C.E., Gates, P.M., Furlong, E.T. and Ferrer, I. (2001). Molecular resolution and fragmentation of fulvic acid by electrospray ionization/multistage tandem mass spectrometry. *Analytical Chemistry*. 73 (7): 1461-1471.

Lek, S. and Guégan, J.F. (1999). Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling*. 120 (2-3): 65-73.

Lek, S., Belaud, A., Baran, P., Dimopoulos, I. and Delacoste, M. (1996). Role of some environmental variables in trout abundance models using neural networks. *Aquatic Living Resources*. 9 (1): 23-29.

Levenberg, K. (1944). A method for the solution of certain nonlinear problems in least squares. *Quarterly Journal of Applied Mathematics*. 2 (2): 164-168.

Lewicka, Z.A., Benedetto, A.F., Benoit, D.N., Yu, W.W., Fortner, J.D. and Colvin, V.L. (2011). The structure, composition, and dimensions of TiO_2 and ZnO nanomaterials in commercial sunscreens. *Nanoparticle Research*. 13 (9): 3607-3617.

Lewis, J. (1996). Turbidity-controlled suspended sediment sampling and load estimation. *Water Resources Research*. 32 (7): 2299-2310.

Li Puma, G., Bono, A., Krishnaiah, D. and Collin, J.G. (2008). Preparation of titanium dioxide photocatalyst loaded onto activated carbon support using chemical vapor deposition: a review paper. *Hazardous Materials*. 157 (2-3): 209-219.

Li, S. and Sun, W. (2011). A comparative study on aggregation / sedimentation of TiO_2 nanoparticles in mono- and binary systems of fulvic acids and Fe (III). *Hazardous Materials*. 197: 70-79.

Lin, D. and Tian, X. (2010). Fate and transport of engineered nanomaterials in the environment. *Environmental Quality*. 39: 1896-1908.

Linsebigler, A.L., Lu, G. and Yates, J.T. (1995). Photocatalysis on TiO_2 surfaces: Principles, mechanisms, and selected results. *Chemical Review*. 95: 735-758.

Liu, H.H. and Cohen, Y. (2014). Multimedia environmental distribution of engineered nanomaterials. *Environmental Science and Technology*. 48 (6): 3281-3292.

Liu, H.H., Surawanvijit, S., Rallo, R., Orkoulas, G. and Cohen, Y. (2011). Analysis of nanoparticle agglomeration in aqueous suspensions via constant-number Monte Carlo simulation. *Environmental science and technology*. 45 (21): 9284-9292.

Liu, R., Ge, Y., Holden, P.. and Cohen, Y. (2014). Visual data exploration of soil bacteria susceptible to engineered nanomaterials. In: *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2014, Bioinformatics and Biomedicine (BIBM), pp. 18-24.

Liu, R., Rallo, R., George, S., Ji, Z., Nair, S., Nel, A.E. and Cohen, Y. (2011). Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. *Small*. 7 (8): 1118-1126.

Liu, X., Chen, G. and Su, C. (2011). Effects of material properties on sedimentation and aggregation of titanium dioxide nanoparticles of anatase and rutile in the aqueous phase. *Colloid and Interface Science*. 363 (1): 84-91.

Liu, X., Tang, K., Harper, B., Steevens, J.A. and Xu, R. (2013). Predictive modeling of nanomaterial exposure effects in biological systems. *International Nanomedicine*. 8 (Suppl 1): 31-43.

Lomer, M.C.E., Thompson, R.P.H., Commisso, J., Keen, L. and Powell, J.J. (2000). Determination of titanium dioxide in foods using inductively coupled plasma optical emission spectrometry. *The Analyst*. 125: 2339-2343.

Lovern, S.B., Strickler, R.J. and Klaper, R. (2007). Behavioral and physiological changes in *Daphnia magna* when exposed to nanoparticle suspensions (titanium dioxide, nano-C_{60}, and $C_{60}H_xC_{70}H_x$). *Environmental Science and Technology*. 41 (12): 4465-4470.

Lövestam, G., Rauscher, H., Roebben, G., Klüttgen, B.S., Gibson, N., Putaud, J. and Stamm, H. (2010). *Considerations on a definition of nanomaterial for regulatory purposes*.

Luttrell, T., Halpegamage, S., Tao, J., Kramer, A., Sutter, E. and Batzill, M. (2014). Why is anatase a better photocatalyst than rutile? Model studies on epitaxial TiO_2 films. *Scientific Reports*. 4 (4043): 1-8.

Ma, H.B., Brennan, A. and Diamond, S.A. (2012). Phototoxicity of TiO_2 nanoparticles under solar radiation to two aquatic species: *Daphnia magna* and Japanese medaka. *Environmental Toxicology and Chemistry*. 31 (7): 1621-1629.

Mackay, D. (1979). Finding fugacity feasible. *Environmental Science and Technology*. 13 (10): 1218-1223.

Mackay, D., Webster, E., Cousins, I., Cahill, T., Foster, K. and Gouin, T. (2001). *An introduction to multimedia models*.

MacLeod, M., Scheringer, M., McKone, T.E. and Hungerbühler, K. (2010). The state of multimedia mass-balance modeling in environmental science and decision-making. *Environmental Science and Technology*. 44. p.p. 8360-8364.

Macwan, D.P., Dave, P.N. and Chaturvedi, S. (2011). A review on nano-TiO_2 sol-gel type syntheses and its applications. *Material Science*. 46 (11): 3669-3686.

Maier, H.R. and Dandy, G.C. (2000). Neural networks for the prediction and forecasting of water resources variables : a review of modelling issues and applications. *Environmental Modelling and Software*. 15 (1): 101-124.

Malvern Instruments (2011). Zeta potential: An Introduction in 30 minutes. *Zetasizer Nano Serles Technical Note. MRK654-01*. 2: 1-6.

Mansfield, C.M., Alloy, M.M., Hamilton, J., Verbeck, G.F., Newton, K., Klaine, S.J. and Roberts, A.P. (2015). Photo-induced toxicity of titanium dioxide nanoparticles to *Daphnia magna* under natural sunlight. *Chemosphere*. 120: 206-210.

March, S.T. and Smith, G.F. (1995). Design and natural science research on information technology. *Decision Support Systems*. 15 (4): 251-266.

Marquardt, D.W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*. 11 (2) ppp. 431-441.

Marsland, S. (2009). *Machine learning: an algorithm perspective*. 1st. Ed. R. Herbrich and T. Graepel (eds.). Boca Raton, New York: CRC Press.

Martyanov, I.N., Savinov, E.N. and Klabunde, K.J. (2003). Influence of solution composition and ultrasonic treatment on optical spectra of TiO_2 aqueous suspensions. *Colloidal Interface Science*. 267 (1): 111-116.

Mash, H.E., Chin, Y., Sigg, L., Hari, R. and Xue, H. (2003). Complexation of copper by zwitterionic aminosulfonic (Good) buffers. *Analytical Chemistry*. 75 (3): 671-677.

Mccleskey, R.B., Nordstrom, D.K. and Ryan, J.N. (2012). Comparison of electrical conductivity calculation methods for natural waters. *Limnology and Oceanography: Methods*. 10: 952-967.

Meesters, J.A.J., Koelmans, A.A., Quik, J.T.K., Hendriks, J.A. and Meent, D. van D. (2014). Multimedia Modeling of Engineered Nanoparticles with SimpleBox4nano: Model definition and evaluation. *Environmental Science and Technology*. 48: 55726-55736.

Mendive, C.B., Bredow, T., Feldhoff, A. and Blesa, M.A. (2009). Adsorption of oxalate on anatase (100) and rutile (110) surfaces in aqueous systems: experimental results vs. theoretical predictions. *Physical Chemistry Chemical Physics*. 11 (100): 1794-1808.

Metzler, D.M., Erdem, A., Tseng, Y.H. and Huang, C.P. (2012). Responses of algal cells to engineered nanoparticles measured as algal cell population, chlorophyll a, and lipid peroxidation: effect of particle size and type. *Nanotechnology*. 237284: 1-12.

Metzler, D.M., Li, M., Erdem, A. and Huang, C.P. (2011). Responses of algae to photocatalytic nano-TiO_2 particles with an emphasis on the effect of particle size. *Chemical Engineering Journal*. 170 (2-3): 538-546.

Meyer, D.E., Curran, M.A. and Gonzalez, M.A. (2009). An examination of existing data for the industrial manufacture and use of nanocomponents and their role in the life cycle impact of nanoproducts. *Environmental Science and Technology*. 43 (5): 1256-1263.

Mihelcic, J.R. and Zimmerman, J.B. (2010). *Environmental Engineering: Fundamentals, Sustainability, Design*. John Wiley and Sons,Inc. USA.

Mitsiadis, T.A., Woloszyk, A. and Jiménez-Rojo, L. (2012). Nanodentistry: combining nanostructured materials and stem cells for dental tissue regeneration. *Nanomedicine*. 7 (11): 1743-1753.

Moller, M.F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*. 6 (4): 525-533.

Money, E.S., Reckhow, K.H. and Wiesner, M.R. (2012). The use of Bayesian networks for nanoparticle risk forecasting: Model formulation and baseline evaluation. *Science of the Total Environment*. 426: 436-445.

Moore, J. (2012). New Zealand's regulation of cosmetic products containing nanomaterials. *Bioethical Inquiry*. 9 (2): 185-188.

Moré, J.J. and Sorensen, D.C. (1983). Computing a trust region step. *SIAM Journal on Scientific and Statistical Computing*. 4 (3): 553-572.

Von Moos, N., Bowen, P. and Slaveykova, V.I. (2014). Bioavailability of inorganic planktonic bacteria and aquatic microalgae in freshwater. *Environmental Science Nano*. 1: 214-232.

Mudunkotuwa, I.A. and Grassian, V.H. (2010). Citric acid adsorption on TiO_2 nanoparticles in aqueous suspensions at acidic and circumneutral pH: surface coverage, surface speciation, and its impact on nanoparticle-nanoparticle interactions. *American Chemical Society*. 132 (42): 14986-14994.

Mueller, N.C. and Nowack, B. (2008). Exposure modeling of engineered nanoparticles in the environment. *Environmental Science and Technology*. 42 (12): 4447-4453.

Murphy, E.M., Zachara, J.M., Smith, S.C., Phillips, J.L. and Wietsma, T.W. (1994). Interaction of hydrophobic organic compounds with mineral-bound humic substances. *Environmental Science and Technology*. 28: 1291-1299.

Murphy, K.P. (2012). *Machine learning: a probabilistic perspective*. 1st. Ed. Cambridge, Massachusetts, USA: The MIT Press.

Musee, N. (2011a). Nanowastes and the environment : Potential new waste management paradigm. *Environmental International*. 37 (1): 112-128.

Musee, N. (2011b). Simulated environmental risk estimation of engineered nanomaterials: A case of cosmetics in Johannesburg City. *Human and Experimental Toxicology*. 30 (9): 1181-1195.

Musee, N., Thwala, M. and Nota, N. (2011). The antibacterial effects of engineered nanomaterials: implications for wastewater treatment plants. *Environmental Monitoring*. 13 (5): 1164-1183.

Mwaanga, P., Carraway, E.R. and Schlautman, M.A. (2014). Preferential sorption of some natural organic matter fractions to titanium dioxide nanoparticles: influence of pH and ionic strength. *Environmental Monitoring and Assessment*. 186 (12): 8833-8844.

Naftaly, U., Intrator, N. and Horn, D. (1997). Optimal ensemble averaging of neural networks. *Network: Computation in Neural Systems*. 8 (3): 283-296.

Nanowerk, L. (2015). *Nano Catalog*. [Online]. 2015. Available from: http://nanowerk.com/nanocatalog/. [Accessed: 11 April 2015].

Nason, J.A., McDowella, S.A. and Callahan, T.W. (2012). Effects of natural organic matter type and concentration on the aggregation of citrate-stabilized gold nanoparticles. *Environmental Monitoring*. 14: 1885-1892.

Navarro, E., Baun, A., Behra, R., Hartmann, N.B., Filser, J., Miao, A.-J., Quigg, A., Santschi, P.H. and Sigg, L. (2008). Environmental behavior and ecotoxicity of engineered nanoparticles to algae, plants, and fungi. *Ecotoxicology*. 17 (5): 372-386.

Neal, C., Jarvie, H., Rowland, P., Lawler, A., Sleep, D. and Scholefield, P. (2011). Titanium in UK rural, agricultural and urban/industrial rivers: geogenic and anthropogenic colloidal/sub-colloidal sources and the significance of within-river retention. *Science of the Total Environment*. 409 (10): 1843-1853.

Necedal, J. and Wright, S.J. (2006). *Numerical optimization*. 2nd Ed. Springer Science + Busines Media, LLC.

Nikinmaa, M. (2014). *An introduction to aquatic toxicology*. 1st Ed. Academic Press.

Nolan, M., Parker, S.C. and Watson, G.W. (2006). Reduction of NO_2 on ceria surfaces. *The Physical Chemistry B*. 110 (5): 2256-2262.

Nota, N.K.C. (2011). *MSc. Thesis: estimated environmental risks of engineered nanomaterials in Gauteng*. Stellenbosch University.

Nowack, B. and Bucheli, T.D. (2007). Occurrence, behavior and effects of nanoparticles in the environment. *Environmental Pollution*. 150 (1): 5-22.

Nowack, B., Krug, H.F. and Height, M. (2011). 120 years of nanosilver History: Implications for policy makers. *Environmental Science and Technology*. 45 (4): 1177-1183.

Nowack, B., Ranville, J.F., Diamond, S., Gallego-Urrea, J.A., Metcalfe, C., Rose, J., Horne, N., Koelmans, A.A. and Klaine, S.J. (2012). Potential scenarios for nanomaterial release and subsequent alteration in the environment. *Environmental Toxicology and Chemistry*. 31 (1): 50-59.

O'Brien, N. and Cummins, E. (2010). Nano-scale pollutants: fate in Irish surface and drinking water regulatory systems. *Human and Ecological Risk Assessment*. 16 (4): 847-872.

Oberdörster, E. (2004). Manufactured nanomaterials (Fullerenes, C_{60}) induce oxidative stress in the brain of juvenile Largemouth Bass. *Environmental Health Perspectives*. 112 (10): 1058-1062.

OECD (2012). *Test No. 211: Daphnia magna Reproduction Test*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789264185203-en.

OECD (2011). *Test No. 201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789264069923-en.

Ohtani, B., Mahaney-Prieto, O.O., Li, D. and Abe, R. (2010). What is Degussa (Evonik) P25? Crystalline composition analysis, reconstruction from isolated pure particles and photocatalytic activity test. *Photochemistry and Photobiology A: Chemistry*. 216: 179-182.

Olden, J.D., Joy, M.K. and Death, R.G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*. 178 (3-4): 389-397.

Ottofuelling, S., Von Der Kammer, F. and Hofmann, T. (2011). Commercial titanium dioxide nanoparticles in both natural and synthetic water: comprehensive multidimensional testing and prediction of aggregation behavior. *Environmental Science and Technology*. 45 (23): 10045-10052.

Palchoudhury, S., Baalousha, M. and Lead, J.R. (2015). Methods for measuring concentration (mass, surface area and number) of nanomaterials. In: M. Baalousha and J. R. Lead (eds.). *Characterization of nanomaterials in complex environmental and biological media*. Elsevier, Frontiers of Nanoscience, 8, pp. 153-181.

Petrova, S.S. and Solov'ev, A.D. (1997). The origin of the method of steepest descent. *Historia Mathematica*. 24 (4): 361-375.

Di Paola, A., Cufalo, G., Addamo, M., Bellardita, M., Campostrini, R., Ischia, M., Ceccato, R. and Palmisano, L. (2008). Photocatalytic activity of nanocrystalline TiO_2 (brookite, rutile and brookite-based) powders prepared by thermohydrolysis of $TiCl_4$ in aqueous chloride solutions. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*. 317 (1-3): 366-376.

Parent, L., Twiss, M.R. and Campbell, P.G.. (1996). Influences of natural dissolved organic matter on the interaction of aluminum with the microalga *Chlorella*: A test of the free-ion model of trace metal toxicity. *Environmental Science and Technology*. 30 (5): 1713-1720.

Parkhust, D.L. and Appelo, C.A.. (1999). *User's guide to PHREEQC (version 2), a computer program for speciation, batch-reaction, one-dimensional transport and inverse geochemical calculations.*

Pathakoti, K., Huang, M., Watts, J.D., He, X. and Hwang, H. (2014). Using experimental data of Escherichia coli to develop a QSAR model for predicting the photo-induced cytotoxicity of metal oxide nanoparticles. *Photochemistry and Photobiology, B: Biology*. 130: 234-240.

Pedersen, F.. and Hansen, P.. (2003). Effects of high pH on the growth and survival of six marine heterotrophic protists. *Marine Ecology Progress Series*. 260: 33-41.

Perego, C., Clemencon, I., Rebours, B., Revel, R., Durupthy, O., Cassaignon, S. and Jolivet, J.-P. (2009). Thermal stability of Brookite - TiO_2 nanoparticles with controlled size and shape: In-situ studies by XRD. In: *Material Research Society Symposium Proceedings*. 2009, pp. 1-7.

Petosa, A.R., Jaisi, D.P., Quevedo, I.R., Elimelech, M. and Tufenkji, N. (2010). Aggregation

and deposition of engineered nanomaterials in aquatic environment: role of physicochemical interactions. *Environmental Science and Technology*. 44 (17): 6532-6549.

Pettibone, J.M., Cwiertny, D.M., Scherer, M. and Grassian, V.H. (2008). Adsorption of organic acids on TiO_2 nanoparticles: effects of pH, nanoparticle size, and nanoparticle aggregation. *Langmuir*. 24 (14): 6659-6667.

Philippe, A. and Schaumann, G.E. (2014). Interactions of dissolved organic matter with natural and engineered inorganic colloids: A review. *Environmental Science and Technology*. 48: 8946-8962.

Piccinno, F., Gottschalk, F., Seeger, S. and Nowack, B. (2012). Industrial production quantities and uses of ten engineered nanomaterials in Europe and the world. *Nanoparticle Research*. 14 (119): 1-11.

Planchon, M., Ferrari, R., Guyot, F., Gélabert, A., Menguy, N., Chanéac, C., Thill, A., Benedetti, M.F. and Spalla, O. (2013). Interaction between *Escherichia coli* and TiO_2 nanoparticles in natural and artificial waters. *Colloids and Surfaces B: Biointerfaces*. 102: 158-164.

Plasmacarb (2003). *Towards the bulk production of fullerenes, nanotubes and carbon black by a clean process*. EU Community Research, p 1-6.

Popov, A.P., Priezzhev, A.V., Lademann, J. and Myllylä, R. (2005). TiO_2 nanoparticles as an effective UV-B radiation skin-protective compound in sunscreens. *Physics D: Applied Physics*. 38 (15): 2564-2570.

Praetorius, A., Scheringer, M. and Hungerbuhler, K. (2012). Development of environmental fate models for engineered nanoparticles-A case study of TiO_2 nanoparticles in the Rhine river. *Environmental Science and Technology*. 46 (12): 6705-6713.

Project on Emerging Nanotechnologies (2015). *Consumer Products Inventory*. [Online]. 2015. Available from: http://www.nanotechproject.org/cpi. [Accessed: 24 April 2015].

Puzyn, T., Leszczynska, D. and Leszczynski, J. (2009). Toward the Development of 'Nano-QSARs': Advances and Challenges. *Small*. 5 (22): 2494-2509.

Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T.P., Michalkova, A., Hwang, H.-M., Toropov, A., Leszczynska, D. and Leszczynski, J. (2011). Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nature Nanotechnology*. 6 (3): 175-178.

Raghavender, A.T., Hong, N.H., Lee, K.J., Jung, M.-H., Skoko, Z., Vasilevskiy, M.,

Cerqueira, M.F. and Samantilleke, A.P. (2013). Nano-ilmenite FeTiO3: synthesis and characterization. *Magnetism and Magnetic Materials*. 331: 129-132.

Rajendran, R., Balakumar, C., Ahammed, M.H.A., Jayakumar, S., Vaideki, K. and Rajesh, E.M. (2010). Use of zinc oxide nanoparticles for production of antimicrobial textiles. *International Engineering, Science and Technology*. 2 (1): 202-208.

Ralph, J. (2012). *The Mindat Mineral and Gem Directory*. 2012. Available from: www.mindat.org/directoy.php. [Accessed: 12 July 2012]

Rand, G.., Wells, P.. and McCarty, L.. (2003). Introduction to aquatic toxicology. In: G. M. Rand (ed.). *Fundamentals of aquatic toxicology: effects, environmental fate, and risk assessment*. Taylor and Francis, pp. 1-67.

Recknagel, F., French, M., Harkonen, P. and Yabunaka, K.-I. (1997). Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling*. 24 (20): 2515-2518.

Reeves, J.F., Davies, S.J., Dodd, N.J.F. and Jha, A.N. (2008). Hydroxyl radicals (• OH) are associated with titanium dioxide (TiO_2) nanoparticle-induced cytotoxicity and oxidative DNA damage in fish cells. *Mutation Research*. 640: 113-122.

Reibold, M., Pätzke, N., Levin, A.A., Kochmann, W., Shakhverdova, I. P., Paufler, P. and Meyer, D.C. (2009). Structure of several historic blades at nanoscale. *Crystal Research and Technology*. 44 (10): 1139-1146.

Reyes-Coronado, D., Rodríguez-Gattorno, G., Espinosa-Pesqueira, M.E., Cab, C., de Coss, R. and Oskam, G. (2008). Phase-pure TiO_2 nanoparticles: anatase, brookite and rutile. *Nanotechnology*. 19 (14): 145605-145614.

Ritz, C. (2010). Toward a unified approach to dose-response modeling in ecotoxicology. *Environmental Toxicology and Chemistry*. 29 (1) ppp. 220-229.

Robichaud, C.O., Uyar, A.L.I.E., Darby, M.R., Zucker, L.G. and Wiesner, M.R. (2009). Estimates of upper bounds and trends in nano-TiO_2 production as a basis for exposure assessment. *Environmental Science and Technology*. 43 (12): 4227-4233.

Roco, M.C., Mirkin, C.A. and Hersam, M.C. (2010). *Nanotechnology research directions for societal needs in 2020*. Springer.

Roebben, G., Rauscher, H., Amenta, V., Aschberger, K., Sanfeliu, A.B., Calzolai, L., Emons, H., Gaillard, C., Gibson, N., Holzwarth, U., Koeber, R., Linsinger, T., Rasmussen, K., Sokull-Klüttgen, B. and Stamm, H. (2014). *Towards a review of the EC recommendation*

for a definition of the term' nanomaterial' Part 2: assessment of collected information concerning the experience with the definition.

Roncaroli, F. and Blesa, M.A. (2011). Kinetics of adsorption of oxalic acid on different titanium dioxide samples. *Journal of Colloid and Interface Science*. 356 (1): 227-233.

Rosenblatt, F. (1962). *Principles of Neurodynamics: perceptrons and the theory of brain mechanisms*. Spartan.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*. 65 (6): 386-408.

Rosi, N.L. and Mirkin, C.A. (2005). Nanostructures in biodiagnostics. *Chemical Reviews*. 105 (4): 1547-1562.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (2006). Learning internal representations by error propagation. In: D. E. Rumelhart, J. L. McClelland, and PDP Research Group (eds.). *Parallel distributed processing: Explorations in the microstructure of cognition*. Massachusetts, USA: The MIT Press.

Sadiq, M.I., Dalai, S., Chandrasekaran, N. and Mukherjee, A. (2011). Ecotoxicity study of titania (TiO_2) NPs on two microalgae species: *Scenedesmus sp.* and Chlorella sp. *Ecotoxicology and Environmental Safety*. 74: 1180-1187.

Sala, O.E., Jackson, R.B., Mooney, H.A. and Howarth, R.W. (2000). *Methods in ecosystem science*. 1st Ed. O. E. Sala, R. B. Jackson, H. A. Mooney, and R. W. Howarth (eds.). New York, USA: Springer Science + Busines Media, LLC.

SANS (2011). *SANS 241 - Specifications for drinking water*. SABS Standards Division.

Scardi, M. and Harding, L.W. (1999). Developing an empirical model of phytoplankton primary production: A neural network case study. *Ecological Modelling*. 120 (2-3): 213-223.

SCENIHR (2007). *Opinion on the scientific aspects of the existing and proposed definitions relating to products of nanoscience and nanotechnologies*. http://ec.europa.eu/health/ph_risk/committees/04_scenihr/docs/scenihr_o_012.pdf.

Schapire, R.E. (1990). The strength of weak learnability. *Machine Learning*. 5 (2): 197-227.

Schapire, R.E. (1999). A brief introduction to boosting. In: *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*. 1999, p. 1401-1406.

Schilling, K., Bradford, B., Castelli, D., Dufour, E., Nash, J.F., Pape, W., Schulte, S., Tooley,

I., van den Bosch, J., Schellauf, F. and Nash, F.J. (2010). Human safety review of 'nano' titanium dioxide and zinc oxide. *Photochemical and Photobiological Sciences : Official European Photochemistry Association and the European Society for Photobiology*. 9 (4): 495-509.

Schmid, K. and Riediker, M. (2008). Use of nanoparticles in Swiss industry: A targeted survey. *Environmental Science and Technology*. 42 (7): 2253-2260.

Seitz, F., Bundschuh, M., Rosenfeldt, R.R. and Schulz, R. (2013). Nanoparticle toxicity in *Daphnia magna* reproduction studies: the importance of test design. *Aquatic Toxicology*. 126: 163-168.

Seitz, F., Luderwald, S., Rosenfeldt, R.R., Schulz, R. and Bundschuh, M. (2015). Aging of TiO_2 nanoparticles transiently increases their toxicity to the pelagic microcrustacean *Daphnia magna*. *PLoS ONE*. 10 (5): 1-14.

Seitz, F., Rosenfeldt, R.R., Schneider, S., Schulz, R. and Bundschuh, M. (2014). Size-, surface- and crystalline structure composition-related effects of titanium dioxide nanoparticles during their aquatic life cycle. *The Science of the Total Environment*. 493: 891-897.

Seo, J., Kim, S., Choi, S., Kwon, D., Yoon, T.H., Kim, W.K., Park, J.W. and Jung, J. (2014). Effects of physiochemical properties of test media on nanoparticle toxicity to *Daphnia magna* straus. *Bulletin of Environmental Contamination and Toxicology*. 93 (3): 257-262.

Shanno, D.F. (1970). Conditioning of Quasi-Newton methods for function minimization. *Mathematics of Computation*. 24 (111): 647-656.

Shapiro, S.S. and Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*. 52 (3/4): 591-611.

Shimpf, M., Caldwell, K.. and Giddings, C.. (eds.) (2000). *Field Flow Fractionation Handbook*. New York, USA: Wiley-Interscience.

Siddiquey, I.A., Ukaji, E., Furusawa, T., Sato, M. and Suzuki, N. (2007). The effects of organic surface treatment by methacryloxypropyltrimethoxysilane on the photostability of TiO_2. *Materials Chemistry and Physics*. 105 (2-3): 162-168.

Sillanpää, M., Paunu, T. and Sainio, P. (2011). Aggregation and deposition of engineered TiO_2 nanoparticles in natural fresh and brackish waters. *Physics: Conference Series 304*. 012018: 1-8.

Site, D.A. (2001). Factors affecting sorption of organic compounds in natural sorbent/water

systems and sorption coefficients for selected pollutants. A review. *Physical and Chemical Reference Data*. 30 (1): 187-439.

Sleighter, R.L. and Hatcher, P.G. (2007). The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *Mass Spectrometry*. 42 (5): 559-574.

Sokołowska, M. and Bal, W. (2005). Cu (II) complexation by '"non-coordinating"'-N-2-hydroxyethylpiperazine- N'-2-ethanesulfonic acid (HEPES buffer). *Inorganic Biochemistry*. 99: 1653-1660.

Solomatine, D., See, L.M. and Abrahart, R.J. (2008). Data-driven modelling: concepts, approaches, and experiences. In: R. J. Abrahart (ed.). *Practical Hydroinformatics*. Berlin Heidelberg: Springer-Verlag, pp. 17-31.

Stemig, A.M., Do, T.A., Yuwomo, V.M., Arnold, W.A. and Penn, L.R. (2014). Goethite nanoparticle aggregation: effects of buffers, metal ions, and 4-chloronitrobenzene reduction. *Environmental Science: Nano*. 1 (5). p.p. 478487.

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*. 36 (2): 111-147.

Stone, V., Nowack, B., Baun, A., Brink, N. Van Den, Kammer, F. Von Der, Dusinska, M., Handy, R., Hankin, S., Hassellöv, M., Joner, E. and Fernandes, T.F. (2010). Nanomaterials for environmental studies: classification, referene material issues, and strategies for physico-chemical characterization. *Science of the Total Environment, The*. 408 (7): 1745-1754.

Sun, D.D. and Lee, P.F. (2012). TiO_2 microsphere for the removal of humic acid from water: Complex surface adsorption mechanisms. *Separation and Purification Technology*. 91: 30-37.

Sun, Y.T., Gottschalk, F., Hungerbühler, K. and Nowack, B. (2014). Comprehensive probabilistic modelling of environmental emissions of engineered nanomaterials. *Environmental Pollution*. 185: 69-76.

Suttiponparnit, K., Jiang, J., Sahu, M., Suvachittanont, S., Charinpanitkul, T. and Biswas, P. (2010). Role of surface area, primary particle size, and crystal phase on titanium dioxide nanoparticle dispersion properties. *Nanoscale Research Letters*. 6 (27): 1-8.

Swingler, K. (1996). *Applying neural networks: A practical guide*. London: Academic Press.

Takeda, K., Suzuki, K., Ishihara, A., Kub-Irie, M., Fu, R., Tabata, M., Oshio, S., Nihei, Y.,

Ihara, T. and Sugamata, M. (2009). Nanoparticles transferred from pregnant mice to their offspring can damage the genital and cranial nerve systems. *Health Science*. 55 (1): 95-102.

Tan, L.-Y., Huang, B., Xu, S., Wei, Z.-B., Yang, L.-Y. and Miao, A.-J. (2016). TiO_2 nanoparticle uptake by the water flea *Daphnia magna* via different routes is calcium-dependent. *Environmental Science and Technology*. 50 (14). p.pp. 7799-7807.

Taniguchi, N. (1974). On the basic concept of 'Nano-Technology'. In: *International Conference in Production Engineering*. 1974, Tokyo, Japan: In Bulletin of the Japan Society of Precision Engineering, pp. 18-23.

Ted, S. (2014). *Environmental risk assessment: A toxicological approach*. 4th Ed. Boca Raton, FL, USA: CRC Press.

The Royal Society and the Royal Academy of Engineering (2004). *Nanoscience and nanotechnologies: opportunities and uncertainties*. London.

Thio, B.J.R., Zhou, D. and Keller, A.A. (2011). Influence of natural organic matter on the aggregation and deposition of titanium dioxide nanoparticles. *Hazardous Materials*. 189 (1-2): 556-563.

Tipping, E. (1981). The adsorption of aquatic humic substances by iron oxides. *Geochimica et Cosmochimica Acta*. 45: 191-199.

Toropov, A. and Leszczynski, J. (2006). A new approach to the characterization of nanomaterials: Predicting Young's modulus by correlation weighting of nanomaterials codes. *Chemical Physics Letters*. 433 (1-3): 125-129.

Toropov, A.A., Leszczynska, D. and Leszczynski, J. (2007). Predicting water solubility and octanol-water partition coefficient for carbon nanotubes based on the chiral vector. *Computational Biology and Chemistry*. 31: 127-128.

Trochim, W.M.K. and Linton, R. (1986). Conceptualization for planning and evaluation. *Evaluation and Program Planning*. 9: 289-308.

UNEP (2007). Emerging challenges - Nanotechnology and the environment. In: *GEO Year Book 2007*. United Nations Environment Programme Division of Early Warning and Assessment, pp. 61-70.

US EPA (2002). *Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms*. U.S. Environmental Protection Agency, Washington DC, EPA/821/R-02/012.

US EPA (2010a). *Nanomaterial case studies: Nanoscale titanium dioxide in water treatment and in topical sunscreen (Final)*. U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-09/057F.

US EPA (2010b). *State of the science literature review: Everything nanosilver*. U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-10/084.

US EPA (2011). *Evaluation guideline for ecological toxicity data in the open literature*. *Procedures for screening, reviewing and using published open literature toxicity data in ecological risk assessments*. U.S. Environmental Protection Agency, Washington, DC, https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/evaluation-guidelines-ecological-toxicity-data-open. [Accessed: 26 September 2014].

Vance, M.E., Kuiken, T., Vejerano, E.P., McGinnis, S.P., Hochella, M.F., Rejeski, D. and Hull, M.S. (2015). Nanotechnology in the real world: Redeveloping the nanomaterial consumer products inventory. *Beilstein Journal of Nanotechnology*. 6: 1769-1780.

Vanloon, G.W. and Duffy, S.J. (2005). Organic matter in water. In: *Environmental chemistry: A global perspective*. Oxford; New York, USA: Oxford University Press, pp. 254-272.

Verwey, E.J.W. and Overbeek, J.T.G. (1948). *Theory of the Stability of Lyophobic Colloids*. Elsevier.

Vighi, M. and Calamari, D. (1993). Prediction of the environmental fate of chemicals. *The Institute Superiore Di Sanita*. 29 (3): 209-223.

Virkutyte, J., Dionysiou, D.D. and Al-Abed, S.R. (2012). Depletion of the protective aluminum hydroxide coating in TiO_2-based sunscreens by swimming pool water ingredients. *Chemical Engineering Journal*. 191: 95-103.

Wang, D., Tejerina, B., Lagzi, I., Kowalczyk, B. and Grzybowski, B.A. (2011). Bridging interactions and selective nanoparticle aggregation mediated by monovalent cations. *ACS Nano*. 5 (1): 530-536.

Wang, M., Liao, L., Zhang, X. and Li, Z. (2012). Adsorption of low concentration humic acid from water by palygorskite. *Applied Clay Science*. 67-68: 164-168.

Warheit, D.B., Hoke, R.A., Finlay, C., Donner, E.M., Reed, K.L. and Sayes, C.M. (2007). Development of a base set of toxicity tests using ultrafine TiO_2 particles as a component of nanoparticle risk management. *Toxicology Letters*. 171: 99-110.

Webb, A.R. (2002). *Statistical Pattern Recognition*. 2nd Ed. Chichester, West Sussex, UK: John Wiley and Sons, Ltd.

Wei, B., Sugiura, N. and Maekawa, T. (2001). Use of artificial neural network in the prediction of algal blooms. *Water Research*. 35 (8): 2022-2028.

Weibel, A., Bouchet, R., Boulc'h, F. and Knauth, P. (2005). The big problem of small particles: a comparison of methods for determination of particle size in nanocrystalline anatase powders. *Chemical Materials*. 17 (10): 2378-2385.

Weir, A., Westerhoff, P., Fabricius, L., Hristovski, K. and Von Goetz, N. (2012). Titanium dioxide nanoparticles in food and personal care products. *Environmental Science and Technology*. 46 (4): 2242-2250.

Wendorf, C.A. (2004). Primer on multiple regression coding: common forms and the additional case of repeated contrasts. *Understanding Statistics*. 3 (1): 47-57.

Westerhoff, P. and Nowack, B. (2013). Searching for global descriptors of engineered nanomaterial fate and transport in the environment. *Accounts of Chemical Research*. 46 (3): 844-853.

Westerhoff, P., Song, G., Hristovski, K. and Kiser, M.A. (2011). Occurrence and removal of titanium at full-scale wastewater treatment plants: implications for TiO_2 nanomaterials. *Environmental Monitoring*. 13 (5): 1195-1203.

WHO (2011). *Guidelines for drinking-water quality*. 4th Ed. Geneva, Switzerland. World Health Organization (WHO). http://www.who.int/water_sanitation_health/publications/2011/dwq_guidelines/en/. [Accessed: 25 March 2015].

Wiench, K., Wohlleben, W., Hisgen, V., Radke, K., Salinas, E., Zok, S. and Landsiedel, R. (2009). Acute and chronic effects of nano-and non-nano-scale TiO_2 and ZnO particles on mobility and reproduction of the freshwater invertebrate *Daphnia magna* . *Chemosphere*. 76 (10): 1356-1365.

Wilcoxon, F. (1945). Individual comparisons of grouped data by ranking methods. *Biometrics Bulletin*. 1 (6): 80-83.

Wiley, B., Sun, Y. and Xia, Y. (2007). Synthesis of silver nanostructures with controlled shapes and properties. Accounts of Chemical Research. 40 (10): 1067-1076.

Winkler, D.A., Burden, F.R., Yan, B., Weissleder, R., Tassa, C., Shaw, S. and Epa, V.C. (2014). Modelling and predicting the biological effects of nanomaterials. SAR and QSAR in Environmental Research. 25: 161-172.

Wolpert, D.H. (1992). Stacked generalization. *Neural Networks*. 5 (2): 241-259.

Wong, S.W.Y., Leung, P.T.Y., Djurisić, A.B. and Leung, K.M.Y. (2010). Toxicities of nano zinc oxide to five marine organisms: influences of aggregate size and ion solubility. *Analytical and Bioanalytical Chemistry*. 396 (2): 609-618.

Wright, D.A. and Welbourn, P. (2009). *Environmental Toxicology*. Cambridge . New York, USA: Cambridge University Press.

Xia, X.R., Monteiro-Riviere, N.A. and Riviere, J.E. (2010). An index for characterization of nanomaterials in biological systems. *Nature Nanotechnology*. 5: 671-675.

Yang, K. and Xing, B. (2010). Adsorption of organic compounds by carbon nanomaterials in aqueous phase: Polanyi theory and its application. *Chemical Reviews*. 110 (10): 5989-6008.

Yang, K., Lin, D. and Xing, B. (2009). Interactions of humic acid with nanosized inorganic oxides. *Langmuir*. 25 (18): 3571-3576.

Yao, J. (2005). Information granulation and granular relationships. In: *2005 IEEE International Conference on Granular Computing*. 2005: 326-329.

Yaser, A.M., Malik, M.I. and Lin, H.T. (2012). Neural Networks. In: *Learning from data*. AMLBook, p. 213.

Yeung, D.S., Cloete, I., Shi, D. and Ng, W.W.Y. (2010). Principles of sensitivity analysis. In: *Sensitivity analysis for neural networks*. Springer Berlin Heidelberg, pp. 17-24.

Young, A.G. and Mcquillan, A.J. (2009). Adsorption / adsorption kinetics from ATR-IR spectroscopy . Aqueous oxalic acid on anatase TiO2. *Langmuir*. 25 (10): 3538-3548.

Zhang, H. and Banfield, J.F. (1998). Thermodynamic analysis of phase stability of nanocrystalline titania. *Materials Chemistry*. 8 (9): 2073-2076.

Zhang, H., Penn, R.L., Hamers, R.J. and Banfield, J.F. (1999). Enhanced adsorption of molecules on surfaces of nanocrystalline particles. *The Journal of Physical Chemistry B*. 103 (22): 4656-4662.

Zhang, R., Zhang, H., Tu, C., Hu, X., Li, L., Luo, Y. and Christie, P. (2015). Facilitated transport of titanium dioxide nanoparticles by humic substances in saturated porous media under acidic conditions. *Journal of Nanoparticle Research*. 17 (4): 1-11.

Zhang, Y., Chen, Y., Westerhoff, P. and Crittenden, J. (2009). Impact of natural organic matter and divalent cations on the stability of aqueous nanoparticles. *Water Research*. 43 (17): 4249-4257.

Zhang, Y., Chen, Y., Westerhoff, P., Hristovski, K. and Crittenden, J.C. (2008). Stability of commercial metal oxide nanoparticles in water. *Water Research*. 42: 2204-2212.

Zhao, J., Idaka, H., Takamura, A., Pelizzetti, E. and Serpone, N. (1993). Photodegradation of surfactants. 11. Zeta potential measurements in the photocatalytic oxidation of surfactants in aqueous TiO_2 dispersions. *Langmuir*. 9 (7): 1646-1650.

Zhou, D., Bennett, S.W. and Keller, A.A. (2012). Increased mobility of metal oxide nanoparticles due to photo and thermal induced disagglomeration. *PloS One*. 7 (5): 1-8.

Zhou, D., Ji, Z., Jiang, X., Dunphy, D.R., Brinker, J. and Keller, A.A. (2013). Influence of material properties on TiO_2 nanoparticle agglomeration. *PLoS One*. 8 (11): 1-8.

Zhu, M., Wang, H., Keller, A.A., Wang, T. and Li, F. (2014). The effect of humic acid on the aggregation of titanium dioxide nanoparticles under different pH and ionic strengths. *Science of the Total Environment*. 487: 375-380.

Zhu, X., Chang, Y. and Chen, Y. (2010). Toxicity and bioaccumulation of TiO_2 nanoparticle aggregates in *Daphnia magna*. *Chemosphere*. 78 (3): 209-215.

Appendices

Appendix 4. 1 Illustrative examples of repositories and data type

Database developers	Inventory data details held	Access status	References
American National Standards Institute's Nanotechnology Standards Panel (ANSI-NSP)	Published documents on frameworks, standards, definitions, techniques and specifications related to nanoscience and nanotoxicology.	Public	http://nanostandards.ansi.org/tiki-index.php?page=DataEntryOutputPub
Danish Consumer Council	Consumer nanoproducts that claim to contain ENMs.	Public	http://nanodb.dk/
Nanomaterial-Biological Interaction (NBI) knowledgebase	Repository of characterized ENMs and their biological interactions.	Pseudo-public	http://nbi.oregonstate.edu/
Nano Health-Environment Commented Database (NHECD)	The impact of nanoparticles on health and environment from scientific articles.	Pseudo-public	http://nhecd.jrc.ec.europa.eu/content/discover-our-intelligent-search
Nano Science and Technology Institute (NSTI)	Information about business, intellectual property, research, and contacts on nanotechnology and industries.	Pseudo-public	http://www.nsti.org/about/database.html
Nanowerk LLC	Companies, Laboratories, nanomaterial and characteristics, nanoproducts and application, training, articles.	Public	http://www.nanowerk.com/nanotechnology_databases.php
Nanomaterial Registry	Registry of curated characterized ENMs based on minimal information about nanomaterials.	Public	https://www.nanomaterialregistry.org/

Appendix 4. 1 Illustrative examples of databases and data type (*Continued*)

Database developers	Inventory data details held	Access status	References
Organization for Economic Corporation and Development (OECD)	Research projects on the safety of manufactured nano-materials, and impact of nanoparticles to health and environment from scientific articles.	Public	http://webnet.oecd.org/NanoMaterials
Project on Emerging Nanotechnologies	Consumer nanoproducts containing nanomaterials, their categories, suppliers, origin, and other analyzed information.	Public	http://www.nanotechproject.org/cpi/
The European Consumers Organization (BEUC) and The European consumer voice in Standardization (ANEC)	Consumer nanoproducts containing nanomaterials analyzed data.	Pseudo-public	www.beuc.org/Content/Default.asp?PageID=2142, MS Excel Files
The Silver Institute	Commercial information and use of silver metal and its nanoproducts	Public	https://www.silverinstitute.org/site/publications/
US EPA's Clean-up information (CLU-IN)	Profiles on nanotechnology-based projects in remediation of chemical pollutants	Public	http://www.clu-in.org/databases/
US EPA-Nanoscale Materials Stewardship Program (NMSP)	Voluntary information on production, characterization, and manufacture of nanoproducts from companies.	Public	http://www.epa.gov/oppt/nano/stewardship.htm

Appendix 4. 2 Summary of data on adsorption of organic matter on $nTiO_2$ in water

Properties and source of $nTiO_2$	Exposure conditions	Study
P25: 17.7 ± 6.1 nm, 50±15 m^2/g[1], anatase (80%)[1], rutile (20%)[1], and uncoated[1]. Evonik Degussa Corp., New Jersey.	**Purified water**[2]**:** pH 5 and 7, 10 mM (KCl, $CaCl_2$), $25°C$. NOM (0.0135 mg/ℓ)[4], HA (1 mg/ℓ). $nTiO_2$ (10 mg/ℓ). 24-hours contact time.	(Chowdhury et al. 2012)
P25: ca. 27±4 nm[3], 50±15 m^2/g[3], anatase (82%)[3], rutile (18%)[3] and uncoated[1]. Evonik Degussa (CAS No. 13463-67-7). **Isotherm model:** Freundlich	**Purified water**: pH 4.8 and 7.8, 10-100 mM (NaCl, KCl, $CaCl_2$, and $MgCl_2$). NOM (0.0135 mg/ℓ)[4]. **Moderately hard standard synthetic water** (EPA 1993): pH 7.9, 4.6 mM, $22°C$. TC (13.73 mg/ℓ)[4] HA, FA and NOM (10 mg/ℓ). $nTiO_2$ (5, 10, 15, 20 and 25 mg/ℓ).48-hours contact time.	(Erhayem and Sohn 2014b)
$nTiO_2$ particles: 4 nm, 219 ± 3 m^2/g, uncoated[6] anatase (100%). Structured and Amorphous Materials Inc. **Isotherm model:** Brunauer-Emmet-Teller (BET)	**Purified water**: pH 2 HCl, pH 4 HCl, pH 6 (25 mM MES buffer), pH 7.5 (25 mM HEPES buffer), 30 mM NaCl, $20°C$. NOM (0.036 mg/ℓ)[4], citric acid (0.1-5 mM). $nTiO_2$ (2,000 mg/ℓ). 24-hours contact time.	(Mudunkotuwa and Grassian 2010)
$nTiO_2$ particles: (i) 5 nm, 219 ± 3 m^2/g[7], and (ii) 32 nm, 41 ± 2 m^2/g, both uncoated[5] anatase. Structured and Amorphous Materials Inc. **$nTiO_2$ particles:** 30 nm rutile (100%), details not provided. **Isotherm model:** Langmuir	**Purified water**: pH 2, 10 mM HCl, 20 mM NaCl, pH 5.5 (25 mM MES buffer), pH 6.5 (25 mM HEPES buffer), $25°C$. NOM (0.036 mg/ℓ)[4] Oxalic and adipic acids (0.5-2.2 mM). $nTiO_2$ (500-6,000 mg/ℓ). 2-hours contact time. **Limitations:** pH 5.5 for adipic acid and pH 6.5 for oxalic acid only.	(Pettibone et al. 2008)
P25: ca. 21 nm, 50±15 m^2/g[3], anatase (82%)[3], rutile (18%)[3] and uncoated[1]. Evonik Degussa Corp., New Jersey.	**Order**: water type, pH, IS (mM), DOC (mg C/ℓ), Ca^{2+} (mM). $nTiO_2$ (25 mg/ℓ). **River water:** 6.89, 1.96, 15.76 and 0.06; **Lake water:** 7.08, 3.46, 14.66 and 0.41; **Groundwater:** 6.53, 14.90, 5.12 and 1.48; **Seawater:** 7.86, 735, 2.65 and 10.3. $25°C$. 30-minutes to 1-hour contact time.	(Chekli et al. 2015)
P25: 21 nm, 50±15 m^2/g[3], anatase (82%)[3], rutile (18%)[3] and uncoated[1]. Evonik Degussa Corp., (CAS No. 13463-67-7). **Isotherm model:** Freundlich	**River water order:** pH, IS (mM), DOC (mg C/ℓ). (i) 7.36, 2.8 and 10.01; (ii) 7.64, 6.7 and 5.33; (iii) 8.04, 18.7 and 3.98; and (iv) 7.96, 342 and 3.21. $25°C$. $nTiO_2$ (12.5 mg/ℓ). 30-minutes to 1-hour contact time.	(Chekli et al. 2015)

Appendix 4. 3 Summary of data collated on aggregation of $nTiO_2$ in aqueous environment

Properties and source of $nTiO_2$	Exposure conditions	Study
NM-103: ca. 20 nm, 60 m^2/g, rutile (100%), alumina and siloxane coats[8]; **NM-104**: ca. 20 nm, 60 m^2/g, rutile (100%), alumina, siloxane, and glycerin coats[8]. **P25**: 21 nm, 50 ± 15 m^2/g, anatase (78%), rutile (14%), amorphous (8%) and uncoated[1]. Evonik Degussa, Germany.	**ASTM synthetic hard water**: pH 7.8, 8.23 mM [4], $20°C$. TC (27.45 mg/ℓ) [5]. $nTiO_2$ (1.0 and 10 mg/ℓ). 24-hours contact time.	(Campos et al. 2013)
P25: 18.4 ± 6.0 nm, anatase (80%)[1], rutile (20%)[1], IEP (6.2) spherical and uncoated[1]. Evonik Degussa Corp. New Jersey.	**Purified water**: pH 5 and 7, 1 and 10 mM (KCl), $25°C$. NOM (0.0135 mg/ℓ) [4]. $nTiO_2$ (100, 200, 400 and 800 mg/ℓ). 100-minutes contact time.	(Chowdhury et al. 2011).
P25: 17.7 ± 6.1 nm, anatase (80%)[1], rutile (20%)[1], and uncoated[1]. Evonik Degussa Corp., New Jersey.	**Purified water**[2]**:** pH 5 and 7, 10 mM (KCl, $CaCl_2$), $25°C$. NOM (0.0135 mg/ℓ)[4], HA (1 mg/ℓ). $nTiO_2$ (10 mg/ℓ). 24-hours contact time.	(Chowdhury et al. 2012)
P25: 25.1±8.2 nm, 51.1 m^2/g, anatase (86%), rutile (14%), uncoated[1]. Evonik Degussa	**Reconstituted moderately hard water**: pH 6.7-7.8, 1.1 mM [9], $22°C$. TC (3.44 mg/ℓ). $nTiO_2$ (0.1, 0.25, 0.5, 200, 800 and 1,500 mg/ℓ). 0-24-hours contact time.	(Ma et al. 2012)
P25: 27 ± 4 nm, 51.5 ± 3 m^2/g, anatase (82%), rutile (18%), semi-spherical, and uncoated[1]. Evonik Degussa Corp. (4168063098).	**Seawater**: pH 8.05, 707 mM, $22 \pm 2°C$. TOC (0.65 mg/ℓ). 15-60 minutes contact time. **Mesocosm freshwater**: pH 8.38, 7.18 mM, $22 \pm 2°C$. TOC (63.45 mg/ℓ). $nTiO_2$ (10, 50 and 100 mg/ℓ). 7-hours contact time.	(Keller et al. 2010)
Hombikat UV100: < 10 nm, 288 m^2/g, anatase (67.2%), amorphous (32.8%), zeta potential (-23 mV) and uncoated. Sachtleben Chemie GmbH. **P25**: ca. 30 nm, 47 m^2/g, anatase (72.6%), rutile (18.4%), amorphous (9%), zeta potential (-21 mV) and uncoated from Frederiksen, Ølgod, Denmark **Hombitan LW-S**: ca. 300 nm, 11.5 m^2/g, anatase (100%), zeta potential (-25 mV) and uncoated. Sachtleben Chemie GmbH.	**Algal test Media**: (ISO 8692, 2002): pH 7.45, 7.55 and 7.6 for Hombikat-UV100, P25, and Hombitan-LW-S $nTiO_2$, respectively, 1.5 mM, 20°C. TC (7.156 mg/ℓ) [5]. $nTiO_2$ (2 mg/ℓ). 0, 2, 6, 48 and 72-hours contact time.	(Hartmann et al. 2010)

Appendix 4. 3 Summary of data collated on aggregation of $nTiO_2$ in aqueous environment (*continued*)

Properties and source of $nTiO_2$	Exposure conditions	Study
P25: 21 nm, 50 m^2/g, anatase (70%)[9], rutile (30%)[9] zeta potential (-32.1 mV) and uncoated. Degussa, Essen, Germany. **A100**: 6 nm, 230 m^2/g, anatase (99%)[9], zeta potential (53.5 mV) and uncoated. Crenox GmbH, Krefeld, Germany.	**ASTM reconstituted hard water**: pH 7.93-8.25, 8.23-8.25 mM, $19.85 - 21.55\,°C$, DOC (8.1 mg/ℓ). 504-hours contact time. $nTiO_2$ (2 mg/ℓ).	(Seitz et al. 2013)
P25: 21 nm[1], 50±15 m^2/g[1], anatase (80%)[1] and rutile (20%)[1] and uncoated[1]. Evonik Degussa GmbH, Germany[1].	**Freshwater (lakes)**: pH 6.1 and 5.0, 0.55 and 0.43 mM[10]. TOC (4.1 and 6.1 mg/ℓ); conductivity (27 and 19 µS/cm), respectively. **Brackish waters (estuarine)**: pH 8.1 and 8.0, 145 and 104 mM[10]. TOC (4.2 and 5.6 mg/ℓ); conductivity (9,780 and 7,040 µS/cm), respectively. $25°C$. $nTiO_2$ (1 and 100 mg/ℓ). 60-minutes contact time.	(Sillanpää et al. 2011)
P25: 19.8 nm, 57 m^2/g, uncoated anatase (88%), rutile (12%), Evonik Degussa, Germany.	**Given order below**: Water type, pH, IS (mM) and DOC (mg/ℓ) and Zeta potential (mV) **EPA very hard water:** 8.4, 9.6, < 0.5 and -10 ± 0.2; **EPA moderately hard water**: 7.4, 9.6, < 0.5 and -16 ± 0.6; **EPA very soft water**: 6.4, 0.3, < 0.5 and -20 ± 1.6; **Seawater:** 8.0, 146.5, < 0.5 and +2.0 ± 2.2; **Tap water:** 8.0, 1.8, 1.03 and -9.0 ± 0.3; **Lake water:** 7.9, 1.6, 2.12 and -15 ± 0.6; **Groundwater:** 7.8, 3.5, 1.57 and -1 ± 0.6; **Peat bog water:** 5.2, 0.5, 37.2, and -26 ± 0.2. $25°C$. $nTiO_2$ (25 mg/ℓ). 15-hours contact time.	(Ottofuelling et al. 2011)
P25: 30 nm, 35-45 m^2/g, uncoated anatase (80%), rutile (20%), non-porous spheres. Evonik Degussa, Germany.	**Purified water**: pH 4, 6, 8, 5 mM (NaCl), Fe^{3+}(0-0.2 mM), $22.5°C$. FA (0-5 mg/ℓ). 24-hours contact time. $nTiO_2$ (50 mg/ℓ).	(Li and Sun, 2011)

Appendix 4. 3 Summary of data collated on aggregation of $nTiO_2$ in aqueous environment (*Continued*)

Properties and source of $nTiO_2$	Exposure conditions	Study
$nTiO_2$ particles: < 40 nm, 50 m^2/g [11], anatase (70%), rutile (30%), uncoated[1]. Sigma Chemicals, St. Louis. MO, USA.	**EPA moderately hard water:** pH 7.6, 4.24 mM [2], $25°C$ [12], TC (13.73 mg/ℓ) [5]. $nTiO_2$ (1.0 and 10 mg/ℓ). 24-hours contact time.	(Kim et al. 2014)
P25: 21 nm, 50±15 m^2/g[3], anatase (82%)[3], rutile (18%)[3], and uncoated[1]. Evonik Degussa Corp., New Jersey. (CAS No. 13463-67-7).	**Given order below**: water type, pH, IS (mM), DOC (mg C/ℓ), Ca^{2+} (mM) **Lake water:** 7.08, 3.46, 14.66 and 0.41. **Seawater:** 7.86, 735, 2.65 and 10.3. **Groundwater:** 6.53, 14.90, 5.12 and 1.48. **River water:** 6.89, 1.96, 15.76 and 0.06. $nTiO_2$ (25 mg/ℓ). $25°C$. 30-minutes to 1-hour contact time.	(Chekli et al. 2015)
Analytical technique: DLS and Flow Field-Flow Fractionation (FIFFF).	**River water order:** pH, IS (mM), DOC (mg C/ℓ). (i) 7.36, 2.8 and 10.01; (ii) 7.64, 6.7 and 5.33; (iii) 8.04, 18.7 and 3.98; and (iv) 7.96, 342 and 3.21. $nTiO_2$ (12.5 mg/ℓ). $25°C$. 30-minutes to 1-hour contact time.	(Chekli et al. 2015)

KEY: (Refers to Appendices 4.2 and 4.3 above, and 4.4 and 4.5 below).

[1] Evonik Degussa Corporation. [2] Chowdhury et al., 2011. [3] Kulacki and Cardinale, 2012. [4] Estimated from background TOC values in purified water. [5] Estimated from values obtained from test protocols. [6] Structured and Amorphous Materials Inc. [7] Mudunkotuwa and Grassian, 2010. [8] Joint Research Centre (JRC) in the European Union (EU) repository. [9] Seitz et al. 2014. [10] Estimated using Lind's Equation (Eqn. 4.8). [11]Fabian et al. 2008. [12]Kim et al., 2010. [13]Takeda et al., 2009. [14]Keller et al. 2010.

Appendix 4. 4 Summary of data collated on toxicity of nTiO2 on algae species

Properties and source of $nTiO_2$	Exposure conditions and test protocols	Study
$nTiO_2$ particles: (i) 140±44 nm (ultrafine agglomerates), 38.5 m^2/g, anatase (21%), rutile (79%), alumina (7%) and silica (1%) coats, and (ii) 380 nm (fine agglomerates), 5.8 m^2/g, rutile (100%), alumina coat (1%). DuPont Inc.	**Synthetic algal assay procedure (AAP)**: pH 7.5 [5], 1.001 mM [5], TC (2.15 mg/ℓ) [5]. $nTiO_2$ (0.1, 1.0, 10 and 100 mg/ℓ). $23.7 - 23.8\,°C$, sunlight (8200-9500 lux), 72-hours contact time. Species: *Pseudokirchneriella subcapitata.* OECD 201 (1984).	(Warheit et al. 2007)
Hombikat UV100: < 10 nm, 288 m^2/g, anatase (67.2%), amorphous (32.8%), zeta potential (-23 mV) and uncoated. Sachtleben Chemie GmbH. **P25**: ca. 30 nm, 47 m^2/g, anatase (72.6%), rutile (18.4%), amorphous (9%), zeta potential (-21 mV) and uncoated. Frederiksen, Ølgod, Denmark **Hombitan LW-S**: ca. 300 nm, 11.5 m^2/g, anatase (100%), zeta potential (-25 mV) and uncoated. Sachtleben Chemie GmbH.	**Algal test Media**: pH 7.5, 7.6 and 7.6 for Hombikat-UV100, P25, and Hombitan-LW-S $nTiO_2$, respectively, 1.66 mM, TC (7.16 mg/ℓ) [5]. $nTiO_2$ (0.6-250 mg/ℓ). $20°C$. Sunlight (80-105 µE/m^2/s). 48 and 72-hours contact time. Species: *P.subcapitata.* Modified ISO 8692 (2002).	(Hartmann et al. 2010)
$nTiO_2$ particles: 25-70 nm, 18.6 m^2/g [13], anatase (100%) [13] and uncoated from Sigma Aldrich.	**Algal test media**: pH 7.75[5], 1.67 mM [5], TC (7.16 mg/ℓ) [5]. $nTiO_2$ (1.0-1,000 mg/ℓ). $24°C$, sunlight (400-500 nm wavelength) and 24, 48 and 72 durations of exposure. *P. subcapitata*. OECD 201 (1984).	(Aruoja et al. 2009)
P25: 27±4 nm, ca. 300 nm (agglomerates), 51.5 m^2/g [14], anatase (82%), rutile (18%) and uncoated[1]. Evonik Degussa Corp. (4168063098).	**Mesocosm freshwater**: pH 8.38, 7.18 mM, TC (63.45 mg/ℓ). $nTiO_2$ (50, 100, 200 and 300 mg/ℓ). $18°C$, sunlight (> 0.002 W/m^2) and ≤ 528-hours contact time. Species: *Chlamydomonas moewusii, Chlorella vulgaris and Scenedesmus quadricauda*.	(Cardinale et al. 2012)
P25: 27±4 nm, 50±15 m^2/g, anatase (82%), rutile (18%) and uncoated[1]. Evonik Degussa Corp., (CAS No. 13463-67-7).	**Mesocosm freshwater**: pH 8.38, 7.18 mM, TC (63.45 mg/ℓ). $nTiO_2$ (50, 100, 200 and 300 mg/ℓ).18 °C, sunlight (1.13 W/m^2) and 24 to 600-hours contact time. Species: *Anabaena spp*., *Navicula subminiscula*, *Nitzschia pusilla*, *Oscillatoria spp*., *Planothidium lanceolatum*, *S. quadricauda, Selenastrum minutum, Spyrogyra communis, Stigeoclonium tenue and Tabularia fasciculate.*	(Kulacki and Cardinale 2012)

Appendix 4. 5 Summary of data collated on toxicity of nTiO2 on *D. magna*

Properties and source of $nTiO_2$	Exposure conditions and test protocols	Study
$nTiO_2$ particles: < 40 nm, <200-800 nm (agglomerates), 48.6 m^2/g [11], anatase (70%), rutile (30%) and uncoated. Sigma Chemicals, St. Louis. MO, USA.	**EPA synthetic hard water**: pH 7.8, 8.16 mM [5], TC (27.46 mg/ℓ) [5]. $25°C$. $nTiO_2$ (0.5, 1, 2.5, 5, 7.5 and 10 mg/ℓ). 48- and 504-hours contact time. EPA (1993).	(Kim et al. 2010)
$nTiO_2$ particles: < 40 nm, 48.6 m^2/g [11], anatase (70%), rutile (30%) and uncoated. Sigma Chemicals, St. Louis. MO, USA.	**EPA moderately hard water**: pH 7.6, 4.24 mM [5], TC (13.73 mg/ℓ) [5]. $25°C$ [12]. $nTiO_2$ (0.5, 1, 2.5, 5, 7.5 and 10 mg/ℓ). 48- and 504-hours contact time. EPA (1993).	(Kim et al. 2014)
P25: 21 nm, 50 m^2/g, anatase (80%), rutile (20%) and uncoated. Degussa, Essen, Germany.	**Reconstituted ISO test media**: pH 8, 8.43 mM [5], TC (9.27 mg/ℓ) [5]. $nTiO_2$ (0.1, 0.5, 1.0 and 5.0 mg/ℓ). $22°C$. 24, 48, 72 and 504-hours contact time. Modified OECD 202 (2004).	(Zhu et al. 2010)
P25: 21 nm, 50 m^2/g[1], anatase (80%)[1], rutile (20%)[1] and uncoated. Degussa, Essen, Germany.	**Elendt 4 media**: pH 8.2 [5], 8.51 mM [5], $20°C$, TC (9.27 mg/ℓ) [5]. $nTiO_2$ (1.19, 1.78, 2.67, 4.0 and 6.0 mg/ℓ). 504-hours contact time. OECD 211 (2008).	(Jacobasch et al. 2014)
NM-103: 20 nm, 60 m^2/g, rutile (100%)[6], alumina and siloxane coats[6]; **NM-104**: 20 nm, 60 m^2/g, rutile (100%) [6], alumina, siloxane, and glycerin coats[6]. **P25**: 21 nm, 50 ± 15 m^2/g, anatase (78%), rutile (14%) and amorphous (8%), uncoated, Evonik Degussa, Germany.	**Synthetic hard water**: pH 7.8, 8.23 mM [5], $20°C$. TC (27.46 mg/ℓ) [5]. $nTiO_2$ (1.0 and 10 mg/ℓ). 504-hours contact time. OECD (1981).	(Campos et al. 2013)
P25: 21 nm, 50 m^2/g, anatase (70%) [9] , rutile (30%) and uncoated. Degussa, Essen, Germany. **A100:** 6 nm, 230 m^2/g, anatase (99%) [9] and uncoated. Crenox GmbH, Krefeld, Germany.	**ASTM reconstituted hard water**: pH 7.93-8.25, 8.23-8.25 mM, $19.85 - 21.55\ °C$, DOC (8.1 mg/ℓ). $nTiO_2$ (0.02, 0.06, -.2, 0.6 and 2 mg/ℓ). 504-hours contact time. OECD 211 (2008).	(Seitz et al. 2013)
$nTiO_2$ particles: (i) 140±44 nm (agglomerates), 38.5 m^2/g, anatase (21%), rutile (79%), Al_2O_3 (7%) and SiO_2 (1%) coats, and (ii) 380 nm (agglomerates), 5.8 m^2/g, rutile (100%), Al_2O_3 coat (1%). DuPont Inc.	**Reconstituted water:** pH 7.5 [5], 8.43 mM [5], TC (9.27 mg/ℓ) [5]. $nTiO_2$ (0.1, 1.0, 10 and 100 mg/ℓ). $20.1 - 20.2°C$, 48-hours contact time. OECD 202 (2004).	(Warheit et al. 2007)

Appendix 4. 6 Categories of data collated from scientific literature on behaviour kinetics and toxicity of ENMs

Category	Broad descriptions of data subcategories
1. Historical information	○ **Sources of information:** detail information on articles, books, company documents, material data sheets, and web links. ○ **Sources of ENMs:** Names of commercial manufacturers, suppliers, brand name, CAS numbers and synthesis methods of pristine ENMs as received. ○ **Exposure methods:** static, semi-static or continuous exposure conditions, renewal or non-renewal of ENMs or test water, renewal rates, acute and chronic toxicity distinguished by time. ○ **Sample preparation and testing methods:** treatment of test water and ENM-for example, filtration, use of buffers and dispersants, dispersion water and mechanisms, volumes of test water, flow rates of test water in continuous systems and agitation mechanisms during tests. ○ **Analytical techniques:** specific methods using the optical probe and lesser scattering, electron and scanning probes, spectroscopic, ion particle/spectrometry probe and thermodynamic techniques in characterizing and measuring responses. ○ **Test protocols:** Existing protocols, for example, ASTM, OECD, and US EPA, laboratory SOPs based on previous similar studies.
2. Static physicochemical properties of $nTiO_2$	○ **Physical and structural properties:** purity, crystal structures and concentration in mixtures, density, solubility, primary or nominal particle size, size distribution in dispersive water, polydispersity index, shape, specific surface area ○ **Surface chemistry properties:** surface coating material, uncoated surfaces, the concentration of surface coating material, zeta potential in dispersive and test water, the point of zero charges, isoelectric point, charge density, and chemical reactivity.
3. Abiotic factors	○ The concentration of ions, micro and macro elements; total dissolved solids, total organic carbon or natural organic matter; added salts (synthetic test water); conductivity, pH, ionic strength, alkalinity, hardness. ○ Temperature, duration of exposure, sunlight exposure cycles and intensity, dissolved oxygen, and concentration of ENMs. (Note: Abiotic categorization of $nTiO_2$ based on potential colloid particle properties.

Appendix 4.6 Categories of data collated from scientific literature on behaviour kinetics and toxicity of $nTiO_2$ engineered nanomaterial (*continued*)

Category	Broad descriptions of data subcategories
4. Biotic factors	○ **Related to biological organisms:** Organism species, age, size and weight, development stage, feeding, population ○ **Related to habitat** (additional data): Mixed or axenic cultures that signify hetero- and homogeneous species exposed, respectively.
5. Behaviour kinetics of $nTiO_2$ in water	○ **Adsorption**: Influence of specific static physicochemical properties of ENMs, and environmental abiotic and biotic factors to adsorption kinetics. Endpoints include adsorption constants, molecules or mass surface coverage of adsorbates on the $nTiO_2$ surface. ○ **Aggregation**: Influence of specific static physicochemical properties of ENMs, and environmental abiotic and biotic factors to aggregation kinetics. Endpoints include hydrodynamic size, fractal sizes, size distribution, diffusion coefficients, and attachment coefficients.
6. Toxicity to algae and *D. magna*	○ Influence of specific static physicochemical properties of ENMs, and environmental abiotic and biotic factors to toxicity. Endpoints measured were many, examples include the following, • **Biological endpoints for algae:** Growth rate (relative Chl. Fluorescence, cells/m ℓ, percent inhibition, Chl-a/day,). Respiration rate (ug O_2/ ℓ. hr.ug chl.-a), biomass growth (percent inhibition or cell density, that is cells/mℓ). • **Biological endpoints for *D. magna*:** Mortality or survival, behavioural changes, for example, immobilization, growth lengths of individuals, number of offspring, number of eggs, fecundity, oxidative stress • **Analogous dose-response measurements**: For example, lethal concentrations(LC_x), or lethal dose(LD_x), where x is percent toxicity effects attributed to $nTiO_2$.