

*This file only contains the abstract to the PhD thesis.  
The full work can be found at <https://biblio.ugent.be/publication/8621716>*

Ir. Chaïm De Mulder

Tying up loose ends:  
optimization of data treatment and  
hydrodynamic model structure of  
the Eindhoven wastewater treatment plant model

Thesis submitted in fulfillment of the requirements for the degree of  
**Doctor (PhD) of Applied Biological Sciences: Environmental Technology**  
Academic year 2018-2019

---

**Rector**

Prof. dr. ir. Rik Van de Walle

**Dean**

Prof. Dr. Ir. Marc Van Meirvenne

**Supervisors****Promoter**

Prof. dr. ir. Ingmar Nopens  
BIOMATH Research group  
Faculty of Bioscience Engineering  
Ghent University, Belgium

**Co-promoter**

Dr. ir. Youri Amerlinck  
BIOMATH research group  
Faculty of Bioscience Engineering  
Ghent University, Belgium

**Examination Committee****Chairman**

Prof. dr. ir. Korneel Rabaey  
CMET research group  
Faculty of Bioscience Engineering  
Ghent University, Belgium

**Secretary**

Prof. dr. Jan Verwaeren  
Kermit Research group  
Faculty of Bioscience Engineering  
Ghent University, Belgium

Dr. ir. Stefan Weijers  
Manager Policy and Innovation  
Waterboard De Dommel  
The Netherlands

Prof. dr. ir. Olivier Potier  
Université de Lorraine  
France

Prof. dr. Ramón Ganigué  
CMET Research group  
Faculty of Bioscience Engineering  
Ghent University, Belgium



*Nederlandse titel:*

Losse eindjes wegwerken: optimalisatie van databehandeling en hydrodynamische modelstructuur van de rioolwaterzuiveringsinstallatie te Eindhoven.

*Please refer to this work as follows:*

Chaïm De Mulder (2019). Tying up loose ends: optimization of data treatment and hydrodynamic model structure of the Eindhoven wastewater treatment plant model. PhD Thesis, Ghent University, Belgium

*Cover images:*

Front: *Reflection, The Basin*, by Ellen Thayer, 2011

Back: xkcd, ([xkcd.com/1599](https://xkcd.com/1599)); license: (<https://creativecommons.org/licenses/by-nc/2.5>)

ISBN: 978-94-6357-223-1

The author and the promoters give permission to use this thesis for consultation and to copy parts of it for personal use only. Every other use is subject to the copyright laws, more specifically the source must be extensively specified when using results from this thesis.



# Contents

<b>Acknowledgments</b>	v
<b>Contents</b>	xiii
<b>List of Abbreviations</b>	xvii
<b>English summary</b>	xxi
<b>Nederlandse samenvatting</b>	xxv

<b>I A PhD on water:</b>	
<b>introduction</b>	1
<b>1 Introduction</b>	3
1.1 The place of water in the world . . . . .	3
The place of water in the world . . . . .	3
1.2 Wastewater treatment: a short history . . . . .	4
Wastewater treatment: a short history . . . . .	4
1.3 Modeling of Wastewater Treatment Plants . . . . .	5
Modeling of Wastewater Treatment Plants . . . . .	5
1.4 The context of this PhD . . . . .	5
The context of this PhD . . . . .	5
1.5 The Wastewater Treatment Plant (WWTP) of Eindhoven . . . . .	6
The WWTP of Eindhoven . . . . .	6
1.5.1 Introduction to the Eindhoven WWTP . . . . .	7
Introduction to the Eindhoven WWTP . . . . .	7

1.5.2 Data at the WWTP of Eindhoven	7
Data at the WWTP of Eindhoven	7
1.5.3 The EHV model	11
The EHV model	11
1.6 Outline and objectives	16
Outline and objectives	16

## II Modeling basics: data and model validation 19

<b>2 Data and modeling: an introduction</b>	<b>21</b>
2.1 Data type, quantity and quality	21
2.2 Data acquisition and reconciliation	23
2.3 Data for model calibration and validation	25
2.4 The Eindhoven context	26
<b>3 Data quality requirements</b>	<b>29</b>
3.1 Introduction	30
3.2 Methodology	30
3.2.1 Impact of sensor calibration on model prediction quality	31
3.2.2 Impact of sensor precision on model prediction quality	32
3.2.3 Simulations	34
3.3 Results	36
3.3.1 Impact of sensor calibration on model prediction quality	36
3.3.2 Impact of sensor precision on model prediction quality	39
3.4 Discussion	44
3.5 Conclusions	46
<b>4 Data Analysis and Reconciliation</b>	<b>49</b>
4.1 Introduction	50
4.2 Methodology	50
4.2.1 Data collection and pretreatment	50
4.2.2 The <i>wvdata</i> package	50
4.2.3 Showcase specifics	57
4.3 Results	60
4.3.1 Data exploration	60
4.3.2 Data filtering	61

4.3.3	Reliability of the filling methods	62
4.3.4	Data filling	63
4.3.5	Dataset size and calculation time	66
4.4	Discussion	67
4.4.1	Application of the <i>wdata</i> package	67
4.4.2	User advantages of the <i>wdata</i> package	70
4.4.3	Scientific relevance of the <i>wdata</i> package	72
4.4.4	Comparison with other software packages	72
4.5	Conclusions	73
<b>5</b>	<b>Model validation</b>	<b>75</b>
5.1	Introduction	76
5.2	Methodology	76
5.2.1	Model and input data	76
5.2.2	Validation data	78
5.2.3	Validation methodology	78
5.3	Results and discussion	80
5.3.1	COD and TSS after the Primary Settling Tank	80
5.3.2	Plant-wide validation	81
5.3.3	Averages and totals	87
5.4	Conclusions	88
<b>III The art of complexity:</b>		
<b>Compartmental Modeling</b>		<b>91</b>
<b>6</b>	<b>Compartmental Modeling: an introduction</b>	<b>93</b>
6.1	The origins of hydrodynamic modeling	93
6.2	Searching the optimal level of complexity: Compartmental Models	95
6.3	Control, sensors and their location at a WWTP	98
6.4	The Eindhoven context	100
<b>7</b>	<b>Compartmental Modeling: Methodology</b>	<b>101</b>
7.1	Introduction	102
7.2	An upgrade in model structure: from CFD to dynamic CM	102
7.2.1	The CFD-ASM-model	104
7.2.2	From CFD to CM	104
7.2.3	CM inclusion in the plant-wide model	105

7.2.4	CFD scenarios	107
7.2.5	Dynamic compartment volumes and exchange flows	109
7.2.6	Implementation	114
7.3	Lessons learned from CM development	115
7.4	The final Compartmental Models	118
7.4.1	Fixed CM	119
7.4.2	Dynamic CM	120
<b>8</b>	<b>Compartmental Modelling: Results</b>	<b>121</b>
8.1	Introduction	122
8.2	Methodology	122
8.2.1	Models and input data	122
8.2.2	Validation data and methodology	123
8.2.3	In-depth model analysis	124
8.3	Results and discussion	125
8.3.1	Validation and model comparison: TIS vs. CM	125
8.3.2	In-depth model analysis	131
8.3.3	Numerics and computation times	136
8.4	Conclusions	137
<b>9</b>	<b>Control optimization using a Compartmental Model</b>	<b>139</b>
9.1	Introduction	140
9.2	Methodology	141
9.2.1	Model and input data	141
9.2.2	Scenario analyses	141
9.2.3	Simulations	145
9.3	Results and discussion	145
9.3.1	Effect of sensor location	145
9.3.2	Current control	147
9.3.3	Correcting for sensor location and the potential of two-sensor control	149
9.3.4	Dry versus wet weather control parameters	153
9.4	Conclusions	153
<b>IV</b>	<b>This is the end:</b>	
	<b>discussion, conclusions and perspectives</b>	<b>157</b>
<b>10</b>	<b>Discussion: the added value of 4 years of research</b>	<b>159</b>

10.1 Classic modeling tools: improvements to the EHV model	160
10.2 A novel tool: the benefits of Compartmental Modeling	162
<b>11 The EHV model: priorities and perspectives</b>	<b>165</b>
11.1 Improved model structure of the EHV model	165
11.1.1 Nitrate prediction and hydrodynamics	166
11.1.2 The three-to-one-lane assumption	167
11.1.3 The added value of a suitable mechanistic settler model	168
11.2 The role of (ww)data at the Eindhoven WWTP and beyond	169
11.3 A digital twin for the Eindhoven WWTP	172
<b>12 Compartmental Modeling: features, flaws and future</b>	<b>175</b>
12.1 The compartmental modeling methodology: where to go next?	175
12.1.1 Number and specifications of CFD scenarios	175
12.1.2 CFD to CM structure	177
12.1.3 Relation between flow rates and compartment volumes/exchange flow rates	178
12.2 Future application of compartmental models	180
<b>Bibliography</b>	<b>184</b>
<b>Appendix</b>	<b>201</b>
<b>Curriculum Vitae</b>	<b>223</b>







# List of Abbreviations

- ANN** Artificial Neural Network.
- ARD** Average Relative Deviation.
- ARIMA** Autoregressive Integrated Moving Average.
- ARMA** Autoregressive Moving Average.
- ASM** Activated Sludge Model.
- ASM1** Activated Sludge Model No 1.
- ASM2d** Activated Sludge Model No 2d.
- AST** Activated Sludge Tank.
- BOD** Biological Oxygen Demand.
- BSM** Benchmark Simulation Model.
- BSM1** Benchmark Activated Sludge Model No 1.
- CAS** Conventional Activated Sludge.
- CFD** Computational Fluid Dynamics.
- CM** Compartmental Model.
- COD** Chemical Oxygen Demand.
- COD<sub>s</sub>** Soluble Chemical Oxygen Demand.

**COD<sub>t</sub>** Total Chemical Oxygen Demand.

**CPD** Change Point Detection.

**CSD** Cumulative Species Distribution.

**CSTR** Continuously Stirred Tank Reactor.

**DO** Dissolved Oxygen.

**EU** European Union.

**GMP** Good Modeling Practice.

**HPC** High Performance Computing.

**IAE** Integral of Absolute Error.

**IAESTE** International Association for the Exchange of Students for Technical Experience.

**IPCC** Intergovernmental Panel on Climate Change.

**IWA** International Water Association.

**LHS** Latin Hypercube Sampling.

**LHSMDU** Latin Hypercube Sampling with multidimensional uniformity.

**LOCF** Last Observation Carried Forward.

**LOWESS** Locally Weighted Scatterplot Smoothing.

**MPC** Model-based Predictive Control.

**MRD** Median Relative Deviation.

**MVS** Multivariate Statistics.

**OED** Optimal Experimental Design.

**P.E.** Person Equivalent.

**PCA** Principal Component Analysis.

**PLC** Programmable Logic Controller.

**PLS** Partial Least Squares.

**PST** Primary Settling Tank.

**RAS** Return Activated Sludge.

**RBT** Rain Buffer Tank.

**RMSE** Root Mean Squared Error.

**RTC** Real-Time Control.

**RTD** Residence Time Distribution.

**SBR** Sequencing Batch Reactor.

**SCADA** Supervisory Control And Data Acquisition.

**SDG** Sustainable Development Goal.

**SSE** Sum of Squared Errors.

**SST** Secondary Settling Tank.

**TIS** Tanks In Series.

**TSS** Total Suspended Solids.

**UCT** University of Cape Town.

**UN** United Nations.

**VFA** Volatile Fatty Acids.

**WEF** World Economic Forum.

**WFD** Water Framework Directive.

**WRRF** Water and Resource Recovery Facility.

**WWTP** Wastewater Treatment Plant.



## English summary

Numerous world-wide organizations and institutes, as well as history itself, incite us to acknowledge the large role water plays and has played in human society. During the latest 200 years, a growing population and an increased degree of industrialization have made clean, potable water an ever more precious resource. Current challenges like urbanization and climate change are likely to further increase the value of clean water for human kind.

This PhD specifically is framed in the context of the treatment of municipal wastewater. Taking in dirty water from the sewer system, Wastewater Treatment Plants (WWTPs) spend energy and other resources to make sure this water is sent to the receiving water bodies in a clean state, to avoid both the spread of diseases and ecosystem damage. Also the WWTP of Eindhoven, The Netherlands, does exactly that.

In the context of the European Water Framework Directive, that demands a good water quality of all European surface waters, Waterboard De Dommel, in charge of the Eindhoven WWTP started a collaborative effort to reach the required quality level in the Dommel river. The collaboration with the BIOMATH research group has over the past 12 years led to the development of a model of the plant, mathematically describing the processes taking place. Although this model has truly proven its value in terms of knowledge acquisition and decision support in the latest years, the involved experts agree that, in order to reach more advanced modeling objectives (e.g. correct dynamic predictions during wet weather events), more incremental improvements to the model are required. This thesis acknowledges this observation and approaches the necessary improvements from two sides, each described in their dedicated part. A first part focuses on providing knowledge and tools that allow to improve the further (incremental) development of the model, the second one describes one of those model structure improvements that are expected to be highly relevant.

Two important features in any modeling context are uncertainty and data analysis and reconciliation. For the former, increased knowledge is crucial when moving further with a model; for the latter, time savings are the most longed-for. Both topics are the main driver for the first two chapters of this work. The uncertainty analysis described in the first chapter allows to link uncertainties in model input data to variations in model predictions. In a more general context, results show that the use of good-quality model input data can be at least as important as a good (kinetic) parameter calibration. The Python package described in the second chapter allows to check for this data quality, fill gaps that originate from the discarding of low-quality data and check the reliability of the resulting dataset. The main advantage of this, especially towards future development, is the time-gain achieved in

the, inherently iterative, modeling exercise. A final chapter in the first part concerns the validation of the Eindhoven model after a necessary update, reflecting operational changes. Although this validation is, in general, satisfying, it does emphasize the need for disruptive model improvements once again.

In the second part of this work, the paradigm of Compartmental Modeling and its potential to yield an improved yet simple hydrodynamic model representation of the Eindhoven **WWTP** is investigated. A methodology that allows to incorporate detailed hydrodynamic data into a relatively simple plant-wide **WWTP** model is developed, described and critically analyzed. The methodology yields good results, although the number of choices to be made while constructing a Compartmental Model (CM) still needs to be narrowed down to end up with a more concise and reproducible work-flow. The increased level of hydrodynamic detail in the Eindhoven model by means of the CM of the outer ring of the Activated Sludge Tanks leads to improved predictions of some key variables at almost equal simulation times. In addition, the new model structure allows to test for the effect of sensor location on plant operation, something that was not possible using the previous model structure. The latter analysis applied to the ammonium-dissolved oxygen control at the plant showed that the current control parameters and sensor locations are close to the optimal ones, although potential for some small improvements was discovered still. The improved prediction quality and broadened model application range brought by the Compartmental Modeling approach are likely to apply to other **WWTP** models as well.

In the concluding chapters, several suggestions for further improvements to the Eindhoven **WWTP** model in general and the **Compartmental Model (CM)** methodology in particular are proposed. Specifically for the Eindhoven case, several structural changes are suggested to increase the prediction quality of the model: (1) application of the **CM** methodology to the middle ring of the Activated Sludge Tanks is expected to improve nitrate predictions, (2) modeling the parallel process lanes separately will increase the level of detail and prove valuable when applying the model in an on-line context and (3) the application of a more accurate settler model for the Secondary Settling Tank will provide good-quality predictions also under wet weather circumstances. On a more general level, the use of the **WWTP** model as an on-line ‘digital twin’ to the real plant, as well as the future role of data and sensors at the plant is discussed.

The **CM** methodology in itself would benefit most from several forms of ‘sensitivity analyses’ quantifying the impact of the choices made during the construction of a Compartmental Model on the model output. In addition to this, the automation of the conversion of **Computational Fluid Dynamics (CFD)** results to a **CM** is seen as crucial. Further application of

CMs will, for the time being, also have to come with a strong promise of valuable knowledge obtained, since the construction of a CM currently still needs to go through the construction of a CFD model. This fact is considered the largest bottleneck preventing a wide application of Compartmental Models.



## Nederlandse samenvatting

Globale organisaties en instellingen, en zelfs de loop van onze eigenste menselijke geschiedenis tonen de grote rol aan die water speelde en nog steeds speelt in onze maatschappij. Tijdens de voorbije 200 jaar hebben een groeiende bevolking en een verhoogde industrialiseringsgraad zuiver en drinkbaar water nog meer tot een kostbaar goed gemaakt. De belangrijkste uitdagingen van de komende eeuw, zoals verstedelijking en klimaatverandering, zullen naar alle waarschijnlijkheid het belang van zuiver water voor de mensheid alleen nog maar doen toenemen.

Dit doctoraat kadert specifiek in de behandeling van huishoudelijk afvalwater. Rioolwaterzuiveringsinstallaties (RWZIs) ontvangen vervuild water via de riolering en gebruiken energie en andere hulpbronnen om ervoor te zorgen dat het water in zuivere staat terug in de ontvangende waterlichamen terecht komt, om zo de verspreiding van ziektes en schade aan ecosystemen te vermijden. Dit is ook precies wat de RWZI van Eindhoven (Nederland), doet.

In de context van de Europese Kaderrichtlijn Water (Eng. Water Framework Directive), die een 'goede' waterkwaliteit eist van all Europese oppervlaktewateren, startte Waterschap De Dommel, de beheerder van de RWZI van Eindhoven, meer dan 10 jaar geleden een samenwerkingsverband op om de gevraagde kwaliteit te bereiken in de rivier de Dommel, die van België naar Nederland stroomt. De samenwerking met de BIOMATH onderzoeksgroep leidde tot de ontwikkeling van een model van de waterzuiveringsinstallatie, dat de processen die plaatsvinden in de reactoren op een wiskundige manier beschrijft. Hoewel dit model reeds zijn waarde bewezen heeft bij het vergaren van kennis en het ondersteunen van (investerings)beslissingen, zijn de betrokken experts het erover eens dat modelaanpassingen van een meer innovatieve aard nodig zijn om tegemoet te komen aan een aantal belangrijke doelstellingen, zoals het correct voorspellen van het effect van regenweer. Dit doctoraat komt aan deze vaststelling tegemoet en benadert de benodigde aanpassingen van twee kanten. In een eerste deel wordt gefocust op het voorzien van kennis en tools die een verdere ontwikkeling van het model kunnen helpen en versnellen; het tweede deel beschrijft en test één van de wijzigingen in modelstructuur die verwacht worden grote verbeteringen met zich mee te brengen.

Twee belangrijke concepten in een modelleringscontext zijn onzekerheid en data-analyse en -preparatie. Voor het eerste is een vergroting van de aanwezige kennis belangrijk, voor het tweede is vooral de mogelijkheid tot tijdswinst een belangrijk aspect. Beide concepten vormen samen de hoofdreden voor de eerste twee hoofdstukken van dit doctoraat. De

onzekerheidsanalyse die beschreven wordt in het eerste hoofdstuk laat toe om onzekerheden op de modelinput te linken aan onzekerheden in de modelvoorspellingen. Op een meer algemeen niveau tonen de resultaten aan dat de kwaliteit van de gebruikte inputdata op zijn minst even belangrijk kan zijn als een goede kalibratie van de (kinetische) modelparameters. Het Python pakket beschreven in het tweede hoofdstuk reikt een mogelijkheid aan om de kwaliteit van precies die inputdata te vrijwaren. Gaten die in een dataset ontstaan doordat de kwaliteit van bepaalde datapunten niet voldoet, kunnen ingevuld worden op basis van verschillende algoritmes, en de betrouwbaarheid van het de resulterende dataset kan geschat worden. Het grote voordeel van het pakket is echter de tijdswinst die geboekt wordt tijdens een modeleerprocedure, die van nature iteratief is, wat de tijdswinst verder vergroot. Een laatste hoofdstuk in het eerste deel behandelt de validatie van het Eindhovenmodel, na het doorvoeren van een update om recente veranderingen in bedrijfsvoering goed in rekening te brengen. Hoewel deze validatie over het algemeen voldoet, wijzen resultaten opnieuw op de nood aan modelupdates van een meer disruptieve aard.

In het tweede deel van dit werk wordt het concept van Compartimentele Modellerings onderzocht, samen met het potentieel ervan om een verbeterde maar nog steeds eenvoudige modelvoorstelling te geven van de reactorhydrodynamica in de Eindhoven RWZI. Een methodologie die toelaat om gedetailleerde hydrodynamische gegevens te includeren in een vrij eenvoudig RWZI-model werd ontwikkeld, beschreven en kritisch geanalyseerd. De bekomen resultaten zijn veelbelovend, al is er nog steeds een groot aantal vrijheidsgraden aanwezig in de procedure. Om een meer gerichte en reproduceerbare methodologie te bekomen, zou dit aantal naar beneden gebracht moeten worden. De toename van hydrodynamisch detail in het Eindhovenmodel zorgt voor verbeterde voorspellingen van een aantal belangrijke variabelen, terwijl de simulatietijd gelijk blijft. Bovendien laat de nieuwe modelstructuur ook toe om het effect van sensorlocatie op de bedrijfsvoering van de installatie te onderzoeken, iets wat met de vorige modelstructuur niet mogelijk was. Zo'n onderzoek, toegepast op de ammonium-zuurstofcontrole aanwezig op de Eindhoven RWZI, toont aan dat de huidige controleparameters en sensorlocaties reeds zeer dicht bij hun optimum zitten, hoewel er nog steeds een beperkt potentieel is voor verbetering. De verhoogde kwaliteit van de voorspellingen and het verbreden van de toepasbaarheid die het Compartimenteel Model (CM) met zich meebrengt, zijn naar alle waarschijnlijkheid ook van toepassing op andere RWZI modellen.

In de concluderende hoofdstukken worden suggesties gegeven voor de verdere ontwikkeling van zowel het Eindhoven model als de CM methodologie. In het specifieke geval van Eindhoven worden de volgende structurele modelveranderingen aangeraden om de kwaliteit van

de modelvoorspellingen verder te verhogen: (1) de toepassing van de CM methodologie op de middenring van de actief slib tanks. De verwachting is dat dit vooral de voorspellingen van nitraat ten goede zal komen. (2) het gescheiden modelleren van de parallelle straten aanwezig op de zuivering zal het detail verhogen en bewijst naar alle waarschijnlijkheid zijn nut bij het gebruik van het model in een online context. (3) het gebruik van een accurater model om de nabezinktanks te simuleren wordt verwacht te leiden tot een verbeterde predictiekwaliteit bij regenweer. Meer algemeen wordt ook het gebruik van het Eindhoven model als een 'digitale tweeling' van de echt installatie en de toekomstige rol van data en sensoren op de zuivering besproken.

De CM methodologie wordt verwacht het meest gebaat te zijn bij een 'gevoeligheidsanalyse' op verschillende vlakken. Deze zou dan de impact van verschillende keuzes gemaakt tijdens de constructie van een Compartimenteel Model op de modeloutput in kaart brengen. Bijkomend wordt de automatisering van het omzetten van CFD-resultaten naar een CM als cruciaal gezien. Een verdere applicatie van CMs zal voorlopig gepaard moeten gaan met een sterke belofte op een verhoogd kennisniveau, omdat het opstellen van een CM tot nader order nog steeds het opstellen van een CFD-model vereist. Dit laatste wordt gezien als het grootste knelpunt voor een wijdverspreide toepassing van Compartimentele Modellen.