Using numerical weather prediction and in-sewer sensor data for realtime monitoring and forecasting in urban drainage-wastewater systems

Pedersen, Jonas Wied

Publication date:
2021

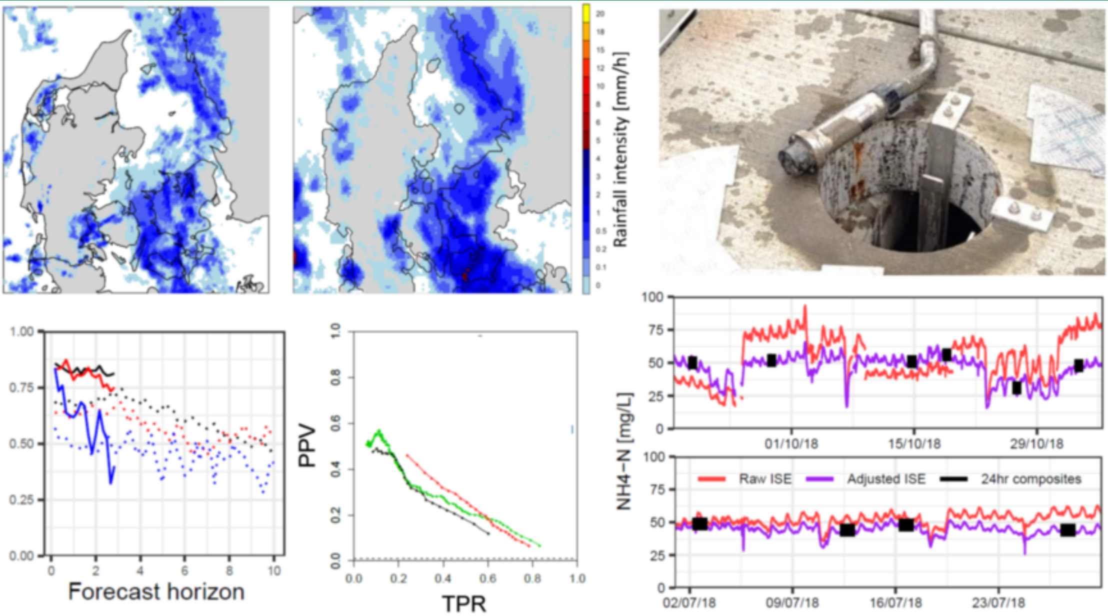
Document Version
Publisher's PDF, also known as Version of record

DTU Environment

Using numerical weather prediction and in-sewer sensor data for real-time monitoring and forecasting in urban drainage-wastewater systems

Jonas Wied Pedersen

PhD Thesis



Using numerical weather prediction and in-sewer sensor data for real-time monitoring and forecasting in urban drainage-wastewater systems

Jonas Wied Pedersen

PhD Thesis
April 2021

DTU Environment
Department of Environmental Engineering
Technical University of Denmark

Using numerical weather prediction and in-sewer sensor data for real-time monitoring and forecasting in urban drainage-wastewater systems

Jonas Wied Pedersen

PhD Thesis, April 2021

The synopsis part of this thesis is available as a pdf-file for download from the DTU research database ORBIT: http://www.orbit.dtu.dk.

Address:	DTU Environment Department of Environmental Engineering Technical University of Denmark Bygningstorvet, Building 115 2800 Kgs. Lyngby Denmark
Phone reception:	+45 4525 1600
Fax:	+45 4593 2850
Homepage:	http://www.env.dtu.dk
E-mail:	reception@env.dtu.dk

Cover: STEP

Preface

This PhD thesis is the culmination of three years of full-time study at the Department of Environmental Engineering, Technical University of Denmark (DTU Environment) during the period October 2016 to January 2021. The period also included temporary leave to contribute to an EU Interreg project and a paternity leave. The PhD project was supervised by Professor Peter Steen Mikkelsen along with co-supervisors Associate Professor Luca Vezzaro, Krüger A/S and DTU Environment, and Professor Henrik Madsen, DTU Compute. The PhD project was a part of the larger Water Smart Cities project funded by Innovation Fund Denmark under the Grand Solution scheme [grant number 5157-00009B].

The thesis is organized in two parts: the first part is a synopsis that provides context and summarizes the main findings of the PhD project; the second part consists of three papers listed below, which have either been published, submitted or are in preparation for peer-reviewed scientific journals. These will be referred to in the text by their paper number written with the Roman numerals **I-III**.

I **Pedersen, J. W.**, Larsen, L. H., Thirsing, C. & Vezzaro, L. (2020). Reconstruction of corrupted datasets from ammonium-ISE sensors at WRRFs through merging with daily composite samples. *Water Research*, 185, 116227. DOI: 10.1016/j.watres.2020.116227.

II **Pedersen, J. W.**, Vezzaro, L., Vedel, H., Thirsing, C., Madsen, H. & Mikkelsen, P. S. (2021). Comparison of high-resolution numerical weather predictions and radar extrapolation forecasts from an urban drainage perspective. Submitted.

III **Pedersen, J. W.**, Courdent, V. A. T., Vezzaro, L., Feddersen, H., Vedel, H., Madsen, H. & Mikkelsen, P. S. (2021). Evaluation of time-lagged numerical weather prediction ensembles for urban runoff forecasting with ROC and PR analysis. Manuscript.

In this online version of the thesis, **Papers I-III** are not included but can be obtained from electronic article databases e.g. via www.orbit.dtu.dk or on request from DTU Environment, Technical University of Denmark, Miljoevej, Building 113, 2800 Kgs. Lyngby, Denmark, info@env.dtu.dk.

The following peer-reviewed journal articles were also prepared during the PhD project but were not part of the thesis:

Vezzaro, L., **Pedersen, J. W.**, Larsen, L. H., Thirsing, C., Duus, L. B. & Mikkelsen, P. S. (2020). Evaluating the performance of a simple phenomenological model for online forecasting of ammonium concentrations at WWTP inlets. *Water Science and Technology*, 81(1), pp. 109-120. DOI: 10.2166/wst.2020.085

Pedersen, A. N., **Pedersen, J. W.**, Vigueras-Rodriguez, A., Brink-Kjær, A., Borup, M. & Mikkelsen, P. S. (2021). The Bellinge data set: Open data and models for community-wide urban drainage systems research. Submitted.

The following conference contributions were also produced during the PhD study:

Vezzaro, L., **Pedersen, J. W.**, Courdent, V. A. T., Löwe, R., & Mikkelsen, P. S. (2017). Towards a domain-based framework for use of rainfall forecasts in control of integrated urban wastewater systems. In Proceedings of 12th IWA Specialized Conference on Instrumentation, Control and Automation, 11-14 June, Québec, Canada, pp. 149-157 (Full paper).

Courdent, V. A. T., **Pedersen, J. W.**, Munk-Nielsen, T., & Mikkelsen, P. S. (2017). Using a time-lagged method to enhance Numerical Weather Prediction for urban drainage applications. In 14th IWA/IAHR International Conference on Urban Drainage, 10-15 September, Prague, Czech Republic, pp. 1639-1642 (Extended abstract).

Pedersen, J. W., Courdent, V. A. T., Vezzaro, L., Vedel, H., Madsen, H., & Mikkelsen, P. S. (2017). Spatial bias and uncertainty in numerical weather predictions for urban runoff forecasts with long time horizons. In 14th IWA/IAHR International Conference on Urban Drainage, 10-15 September, Prague, Czech Republic, pp. 168-171 (Extended abstract).

Pedersen, J. W., Courdent, V. A. T., Vezzaro, L., Madsen, H., & Mikkelsen, P. S. (2017). Urban runoff forecasting with ensemble weather predictions. In 15th Nordic Wastewater Conference, 10-12 October, Aarhus, Denmark, 2 pp (Abstract).

Pedersen, J. W., Vezzaro, L., Vedel, H., Madsen, H., & Mikkelsen, P. S. (2018). Performance of High-Resolution Numerical Weather Predictions with a Rapid Updating Cycle for Urban Runoff Forecasting. In 11th International Conference on Urban Drainage Modelling, 23-26 September, Palermo, Italy. pp. 438-441 (Extended abstract).

Pedersen, J. W., Vezzaro, L., Madsen, H., & Mikkelsen, P. S. (2018). Ensemble forecasts of urban runoff from a deterministic Numerical Weather Prediction model by use of spatial neighborhood sampling. In Rainfall Monitoring, Modelling and Forecasting in Urban Environment. UrbanRain18: 11th International Workshop on Precipitation in Urban Areas, 5-7 December, St. Moritz, Switzerland, pp. 89-91 (Extended abstract).

Stentoft, P. A., Vezzaro, L., Courdent, V., **Pedersen, J. W.**, Thomsen, H. A., Mikkelsen, P. S., Tisserand, B. & Amiel, C. (2019). Real Time Forecasting of Flows and Loads to WWTPs for Enhanced Hydraulic and Biological Capacity during Stormwater Events. In 10th edition of the NOVATECH conference, 2-4 July, Lyon, France, 4 pp (Extended abstract).

Vezzaro, L., **Pedersen, J. W.**, Larsen, L. H., Thirsing, C., Duus, L. B., Breinholt, A., & Mikkelsen, P. S. (2019). Evaluating the performance of a simple phenomenological model for online forecasting of ammonium concentrations. In 9th International Conference on Sewer Processes & Networks, 27-30 August, Aalborg, Denmark, 13 pp (Full paper).

Vezzaro, L., **Pedersen, J. W.**, Larsen, L. H., & Thirsing, C. (2020). Online forecasting of flows and ammonia load at WWTP inlet. In 14th annual Water Research Conference: Danish Water Forum (DWF), Copenhagen, Denmark, pp. 37 (Abstract).

Pedersen, A. N., **Pedersen, J. W.**, Borup, M., Brink-Kjær, Christensen, L. E. & Mikkelsen, P.S. (2021). Use of signatures for systematic diagnostic comparison of time series from urban drainage models and data. Accepted for presentation at the 15th IWA/IAHR International Conference on Urban Drainage, September 2021, Melbourne, Australia (Abstract, postponed due to the Covid-19 pandemic).

Pedersen, J. W., Vezzaro, L., Vedel, H., Madsen, H., Mikkelsen, P. S. (2021): Numerical weather predictions (NWP) as a new source of information for improving the operation of urban drainage and wastewater system. IWA WWC&E, 9-14 May 2021 (Abstract, postponed to September 2022 due to the Covid-19 pandemic).

Acknowledgements

I owe a big thank you to my PhD supervisors. To Peter and Henrik for inspiring supervisor meetings, which I always walked away from thinking that my work could change the world. A million ideas arise when the two of you are in the same room. Thanks to Luca for continually reminding me that a PhD student does not need to follow up on all of his supervisors' ideas, for helping me get things finished, and for our many one-on-one talks. Thanks to Peter for being so warm-hearted and helpful when my personal life got in the way of a simple, linear three-year PhD project.

Thank you to my colleagues at DTU Environment, especially to the group of PhD students in the urban water systems section. You have made it a fun and engaging place to work. A special thanks to the Miljø10 people who have been with me since my first academic baby steps at DTU more than 10 years ago. I am proud to have you as some of my very best friends.

This PhD project was carried out as part of the Water Smart Cities project funded by Innovation Fund Denmark. I would like to thank the project partners for inspiring conversations, workshops and collaborations. Especially Carsten Thirsing and Henrik Vedel who took a keen interest in my work. Also thanks to Krüger for welcoming me at their office and in their interesting project discussions. Thanks to Vianney for helping me get a good start to working with NWP data.

I had the honor of co-supervising some talented, young BSc and MSc thesis students. Some of which developed ideas related to the PhD project that would later go on to inform scientific publications.

Thanks to Jörg Rieckermann and the rest of Eawag for hosting me on a research stay for four months. A special thanks to the PhD group for making me feel like a genuine part of their team.

Elers Kollegium was my home during most of the PhD project. I don't have words to express how thankful I am for that. The atmosphere, history, traditions, and the people I lived there with made it such a wonderful time.

I would also like to thank my family, and my parents-in-law for providing a safe COVID-19 refuge and thesis writing boot camp on their country-side farm. Finally, there is my dear Anne. I love you, and I am immensely grateful to have you and little Erik in my life.

Summary

Urban drainage and wastewater systems are responsible for protecting the environment against pollution and the public against diseases and flooding. These systems have traditionally been engineered as static solutions but the current wave of digitalization means that they are transitioning into actively managed assets. Real-time operations aim to accurately monitor the current state of the system, forecast its near future behavior, and based on this control actuators that allow for flexible performance. These efforts are often built on advanced algorithms that require high-quality input data to properly function. Since rainfall and ammonium concentrations in wastewater are some of the most important variables for general system performance, this thesis deals with obtaining good data on these two aspects. The main research objectives are about how to improve in-sewer measurements of ammonium with ammonium ion-selective electrodes (A-ISE), and how to use numerical weather prediction (NWP) for forecasting rainfall and flow in sewers.

Wastewater from households contain ammonium, which can have serious detrimental environmental effects if discharged into surface waters. It is therefore important that water resource recovery facilities (WRRFs) can accurately monitor, forecast, and ultimately remove it from the wastewater they receive. A-ISE technology has the advantage of measuring directly in the wastewater stream while being cheap to purchase and operate. However, it is also generally regarded as an unreliable data source prone to several types of errors. A one-year measurement campaign at a WRRF highlighted that the currently recommend approach to A-ISE sensor recalibration based on grab samples is inadequate. The result was a raw signal with erratic jumps and effects of drifting. A methodology to correct the errors in the signal was therefore developed based on integrating information from the A-ISE sensors and 24-h volume-proportional composite samples. The composite samples are available at many WRRFs and the methodology can thus be used without additional operational costs. The corrected signal provided a much more reliable estimate of ammonium concentrations, and could be used to estimate software sensors with more precise predictions. While there are still improvements to be made within use of A-ISE for monitoring ammonium in wastewater and to the developed methodology, the thesis has made major progress towards a measurement setup that can deliver reliable A-ISE data to wastewater managers.

NWP predicts rainfall through large-scale simulations of atmospheric physics and is the main alternative to radar extrapolation forecasts, which are more commonly used for urban drainage applications. However, the collective experiences with NWP for urban drainage purposes are still rather few. The thesis therefore reviewed these experiences and extracted key lessons for how to use it well, and further investigated use cases for two different NWP products. Previous research into NWP use for urban drainage issues was grouped into four main topics: (1) generic rain and flow forecasting, (2) urban pluvial flood forecasting, (3) real-time control, and (4) post-processing. Based on this, advice were given on how to make sure that the scope and resolutions of a chosen NWP product, hydrological model, and decision algorithm are fit for the purpose they are intended to fulfil.

In general, it is an issue that many published studies have been built on small samples of a few rain events, which often leads to inconclusive results. This thesis investigated NWP performance with a large forecast archive of more than 100 rain events, which quantified how forecast performance was dependent on the type of weather event. Dynamic events with a high degree of evolution over time and events that consisted of small and scattered rain cells were difficult to predict. The NWP product could successfully be used to control a wet weather switch at a WRRF, which led to improved performance compared to a reactive control setup based on real-time rain gauge measurements.

An intuitive and easy-to-implement post-processing method based on time-lagging was used to enhance a NWP ensemble product. The method was able to use information on forecast consistency from consecutive forecasts, and was used to make sewer flow predictions. Time-lagged forecasts were able to compete with a more well-known post-processing method based on spatial neighborhoods.

NWP is becoming available as an open data source in many countries, and improvements in data resolutions and assimilation techniques are making it increasingly attractive for urban water purposes. With the review of how NWP has been used in the past and the strides made towards using these data for predictions and decision-making, the thesis aims to increase the uptake of NWP for real-time operations in urban drainage and wastewater systems.

Dansk sammenfatning

Afløbs- og spildevandssystemer er ansvarlige for at beskytte miljøet mod forurening og befolkningen mod sygdomme samt oversvømmelser. Systemerne har traditionelt været konstrueret som statiske løsninger, men den igangværende bølge af digitalisering i samfundet betyder, at systemerne er ved at udvikle sig til at blive dynamiske aktiver. God realtidshåndtering af systemerne består af at kunne overvåge deres nuværende status, forudsige hvordan de opfører sig i den nærmeste fremtid og baseret på dette styre aktuatorer, der muliggør fleksibel ydeevne. Alt dette er ofte baseret på avancerede algoritmer, der kræver data af høj kvalitet for at fungere ordentligt. Da regnmængder og ammoniumskoncentrationer i spildevand er nogle af de vigtigste variabler at have styr på, vil denne afhandling handle om at skaffe gode data om disse to aspekter. Afhandlingens centrale forskningsspørgsmål omhandler, hvordan man kan forbedre målinger af ammonium med ion-selektive elektroder (A-ISE), og hvordan man kan bruge numeriske vejrprognoser (NVP) til at forudsige regn og vandmængder i kloakker.

Spildevand fra husholdninger indeholder ammonium, som kan medføre seriøse negative påvirkninger, hvis det udledes til vandmiljøet. Det er derfor vigtigt, at spildevandsrenseanlæg præcist kan overvåge, forudsige og i sidste ende fjerne ammonium fra det spildevand, de modtager. A-ISE sensorer har den fordel, at de kan nedsættes direkte i det rå spildevand, samtidigt med at de er billige at købe og anvende. Teknologien anses dog generelt også for at være upålidelig datakilde, som lider under flere forskellige typer af fejl. En målekampagne på et års længde viste, at den nuværende anbefalede måde at genkalibrere A-ISE sensorer på, som er baseret på håndholdte prøver, er utilstrækkelig. Resultatet af målekampagnen var rådata, der både drev og indeholdte uregelmæssige hop. I afhandlingen er der derfor udviklet en metode, som kan korrigere for disse fejl ved at integrere målinger fra A-ISE sensorerne og volumen-proportionale døgnprøver. Metoden kan implementeres uden yderligere omkostninger, da døgnprøverne allerede udtages på mange renseanlæg. De korrigerede data gav et meget bedre estimat af ammoniumskoncentrationerne og kunne bruges til at forbedre præcisionen af en software-sensor. Selvom både den udviklede metode og de generelle vejledninger til brug af A-ISE teknologi stadigt kan forbedres, så har afhandlingen gjort store fremskridt i forhold til at udvikle et målesystem, som kan forsyne brugere med pålidelige A-ISE data til spildevandshåndtering.

NVP forudsiger regn gennem simuleringer af atmosfærefysik på stor skala, og de er det primære alternativ til radar-baserede ekstrapolationsprognoser, som er bredere anvendt i afløbsbranchen. Branchens erfaring med NVP er stadig lille og spredt. Afhandlingen forsøger derfor at skabe et samlet overblik over disse samt at destillere nogle vigtige læringspunkter for god brug af denne datatype. Den nuværende forskning i NVP for afløbsorienterede formål kan opdeles i fire emner: (1) generiske forudsigelser af regn og flow, (2) forudsigelser af regnbetingede oversvømmelser i byer, (3) realtidsstyring og (4) post-processering. Derudover har forskningen i denne afhandling også yderligere undersøgt to forskellige formål for brug af NVP data. Baseret på dette udstikkes der retningslinjer for, hvordan man kombinerer NVP dataprodukter, hydrologiske modeller og beslutningsalgoritmer på en hensigtsmæssig måde.

Det er et generelt problem, at meget udgivet forskning er baseret på små analyser med en håndfuld regnhændelser, hvilket ofte leder til vage og ufuldstændige konklusioner. For at modvirke dette er resultaterne i denne afhandling baseret på store arkiver af historiske prognoser med mere end 100 renghændelser. Dette har bl.a. muliggjort en kvantificering af, hvordan forskellige vejrtyper påvirker prognosernes nøjagtighed. Dynamiske hændelser, der udvikler sig meget over tid, samt hændelser bestående af små og spredte byger var sværest at forudsige. Afhandlingen har også vist, at et NVP produkt succesfuldt kan anvendes til at styre, hvornår et renseanlæg skal skifte mellem tør- og regnvejrsoptimeret styring.

En intuitiv metode til post-processering af NVP ensembledata, der er nem at implementere, er også blevet vurderet i afhandlingen. Metoden kunne udnytte information, om hvor konsistente efterfølgende prognoser er, og blev brugt til at forudsige afstrømning af vand i kloakker. Denne ”time-lag”-metode viste sig at være konkurrencedygtig sammenlignet med en mere velkendt post-processeringsalgoritme baseret på rumlige naboområder.

NVP er på vej til at blive en åben og gratis datakilde i mange lande, og vedvarende forbedringer i dataopløsning og dataassimileringsmetoder gør dem mere og mere attraktive for vandbranchen. Med det givne overblik over tidligere brug af NVP, de udstukne retningslinjer samt de præsenterede dataanalyser sigter afhandlingen mod at øge brugen af NVP for realtidshåndtering af både afløbs- og spildevandssystemer.

Table of contents

Abbreviations

A-ISE	Ammonium ion-selective electrode
CSO	Combined sewer overflow
DA	Data assimilation
DMI	Danish meteorological institute
ECMWF	European centre for medium-range weather forecasts
EPS	Ensemble prediction system
FPR	False positives rate
LAM	Limited area model
MPC	Model predictive control
NCEP	National centers for environmental prediction (US)
NH_4^+	Ammonium
NWP	Numerical weather prediction
PR	Precision-recall
ROC	Relative operating characteristics
RTC	Real-time control
SOP	Standard operating procedure
TPR	True positives rate
WRRF	Water resource recovery facility
WWTP	Wastewater treatment plant

1 Introduction and background

1.1 Traditional urban drainage and wastewater systems

Urban drainage and wastewater systems are responsible for safely managing the stormwater runoff caused by rainfall over cities and the wastewater produced by households and industries. These functions are so important for society that they have been named one of the greatest medical advances in human history together with clean water supply (Ferriman, 2007). Drainage systems are generally split into two types of systems: combined sewers where storm- and wastewater are conveyed in the same pipes, and separate sewers where the two never mix. Many cities are dominated by centralized drainage and wastewater solutions, where the combined water and the wastewater component of separate systems are transported to a water resources recovery facility (WRRF). Here, the water is safely treated before it is discharged into a recipient water body.

Installation of centralized systems require massive infrastructure investments, and they are thus typically planned to have lifespans of many decades or even a century. The infrastructure present in many cities today is therefore old and have to deal with growing pressures that were not accounted for by their designers (Neumann et al., 2015). Some of the most pertinent ones are large-scale urbanization, changing rainfall patterns due to climate change, and more stringent regulations on operational costs and environmental impacts.

1.2 The digital revolution is here

At the same time, the ongoing revolution in information and communications technology in society at large is disrupting how the water sector functions. A plethora of new concepts and their relevance for the water sector have emerged in recent years such as digitalization, Smart Cities (Albino et al., 2015), Water 4.0 (Sedlak, 2014), Internet of Things (Atzori et al., 2010), artificial intelligence and big data (Garrido-Baserba et al., 2020). All of these concepts point towards the same key trends:

- Urban water systems are transforming from static installations into dynamic, responsive systems that react to changing needs and conditions in space and time (García et al., 2015; Kerkez et al., 2016).

- The physical systems are becoming cyber-physical with digital platforms, sensor data, modeling techniques, and communications technology playing an increasingly central role in planning and operations (Blumensaat et al., 2019; Eggimann et al., 2017; Kerkez et al., 2016).
- They are expected to deal with an increasingly diverse set of societal needs through integration with other systems. Examples are extraction and recycling of valuable nutrients from wastewater, supplying cities with district heating, and for stormwater infrastructure to provide co-benefits such as additional value to public health (Alves et al., 2018; Grant et al., 2012; van Loosdrecht and Brdjanovic, 2014).

Digital platforms that integrate sensor data and advanced modeling techniques can be used to create so-called "digital twins" of the physical systems (Autiosalo et al., 2019; Wright and Davidson, 2020). These will allow improved system understanding especially for locations that are not directly monitored (Haimi et al., 2013), better performance assessment and reporting, as well as aid with long-term planning through simulations of potential future scenarios (Löwe et al., 2017; Rauch et al., 2017). For urban drainage systems, they may allow investigations of sewer condition analysis (Laakso et al., 2018), groundwater infiltration (Karpf and Krebs, 2011), monitoring of combined sewer overflows (Zhang et al., 2018), etc. For WRRFs, they may assist in analysis and modeling of wastewater inflow composition (Martin and Vanrolleghem, 2014), assessment of energy efficiency (Panepinto et al., 2016), plant-wide control (Solon et al., 2017), etc.

1.3 Real-time operations

Real-time operations of urban water infrastructure consist of three main components: (1) monitoring the present state of the system, (2) forecasting future states, and based on these (3) taking action, e.g. through issuing warnings or controlling system actuators for flexible functioning.

Real-time monitoring of urban water infrastructure is becoming increasingly feasible with decreases in the cost of installation and maintenance of many types of online sensors. These trends are leading towards so-called "ubiquitous sensing" where water infrastructure will be systematically monitored in many locations (Blumensaat et al., 2019; Hill et al., 2014). In addition to physical sensors, real-time models can also produce estimates of the current system state, which is known in the wastewater literature as "software" or “soft” sensors (Haimi et al., 2013). The current increase in data gathering is so strong that a recent horizon scan survey concluded that one of the key future priorities of water managers should be avoiding the risk of drowning in data (Blumensaat et al., 2019). Therrien et al. (2020) have therefore outlined the steps that have to be conducted successfully before raw data becomes useful for system comprehension, modeling, and actions: (1) proper data collection, (2) pre-processing, (3) storage, and (4) mining for patterns.

Forecasting is often done with real-time models of which there are many different kinds. There are several different “spectra” that characterize real-time models. They range from physically-based to data-driven in terms of the amount of physics they incorporate, and from white-box to black-box in terms of how directly interpretable their internal states and computations are. They also range from distributed to lumped in their spatial aggregation, and from deterministic to stochastic in whether they include random processes. Common for many forecasting models are that they require predicted values of their inputs and continual updating of initial conditions through data assimilation (Hutton et al., 2014; Lund et al., 2019; Pedersen et al., 2016).

Control of the systems can be “passive” based on static rules for actuator settings, or “active” through real-time control (RTC) algorithms that adapt to changing conditions (García et al., 2015; Lund et al., 2018). RTC algorithms may be based solely on real-time observations (reactive) or also on forecasts (predictive), they can be manually operated or automatic, and their scope can range from single subcomponents (local) to system-wide management (global) (Lund et al., 2018). Finally, the often-used term model predictive control (MPC) describes setups that use models to simulate potential future system trajectories and choose the optimal course of action (García et al., 2015).

1.4 The importance of high-quality input data

While many advanced forecasting and control schemes have been developed in the scientific literature, the transition from academic desktops to real-life implementation has been less successful. Lund et al. (2018) reviewed the MPC literature for urban drainage applications and concluded that very few publications contain actual case implementations. Most studies simply use synthetic rainfall data or “perfect forecasts” using historical observations as the forecasted values. Only a single study had applied actual rainfall forecasts to evaluate their MPC algorithm (Löwe et al., 2016). The MPC literature has thus almost exclusively focused on developing control algorithms without considering the rainfall inputs that are going to feed them (Lund et al., 2018).

There is large need for working on good inputs for these algorithms in the form of high-quality sensor data and forecasts of boundary conditions (Kerkez et al., 2016). Otherwise it is an open question how many of our advanced algorithms that will be robust enough to leave their perfect or synthetic inputs behind and face the uncertainties and errors that characterize real operational inputs.

1.5 Ammonium monitoring and rainfall predictions

This thesis investigates input data that relate to two central variables in urban drainage and wastewater management: ammonium (NH_4^+) and rainfall. NH_4^+ is one of the main pollutants in wastewater and mostly originates from urine. It may promote eutrophication if it is discharged to the environment, which in turn can cause oxygen depletion in water bodies. At high pH values ammonium turns into ammonia (NH_3), which is toxic to aquatic organisms. A large part of the operations at modern WRRFs are dedicated to removal and reuse of nitrogen-containing compounds including NH_4^+. Knowledge of NH_4^+ concentrations is especially important for operating the aeration in the biological step of WRRFs (Åmand and Carlsson, 2012; Rieger et al., 2014). In combined sewer systems, rainfall is the source of the stormwater component and creates major variations in the hydraulic loading. In large quantities it causes issues such as bypass of untreated or partly treated wastewater, combined sewer overflow (CSO), and even flooding. In smaller quantities, rainfall increases the hydraulic load at WRRFs and lowers their treatment performance. Good forecasts of rainfall may e.g. be used to control in-sewer storage tanks to manage the hydraulic loading in space and time (Löwe et al., 2016).

The specific types of data that are examined in this thesis are ammonium ion-selective electrodes (A-ISE) for monitoring of NH_4^+, and numerical weather prediction (NWP) for rainfall and flow forecasting. Each data source have properties that are highly useful to urban water management. NWP simulates the physical processes of atmospheric motion. It is therefore able to produce rainfall forecasts with forecast horizons that exceed those of the more commonly used radar-based “nowcasts”, which are made with extrapolation of spatial weather radar data. A-ISE sensors are relatively cheap, can be placed directly in the raw wastewater, and produce data at fine temporal resolution.

However, A-ISE sensors have a reputation for being difficult to maintain and producing unreliable and drifting signals, while rainfall forecasts from NWP are regarded as highly uncertain and of poor spatiotemporal resolution.

1.6 Research objectives

The hypothesis of this thesis is that both low-cost in-sewer sensors targeting ammonium and NWP can be highly useful for real-time operations at urban water utilities. The thesis will evaluate the current maintenance protocols of A-ISE sensors and aim to develop modeling techniques that increase signal reliability without additional costs of operation. It will also review and assess the limited number of published NWP applications for urban drainage purposes, provide its own assessment of a promising NWP product, and investigate a technique for enhancing raw NWP output.

The thesis will specifically address the following research questions:

1. How can we recover and reconstruct a useful signal from A-ISE sensors in wastewater applications when there are serious data quality issues, and is it possible to do so without additional costs of operation?
2. Which experiences does the urban drainage community have with NWP, which recommendations for best practices can be distilled from these, and where are future developments needed?
3. How do rainfall forecasts from NWP perform at the small spatial scales of urban drainage and wastewater management, and how do they compare with standard, well-known radar nowcasts?
4. How can a simple post-processing method such as time-lagging enhance the use of NWP ensembles?

1.7 Thesis outline

The first research question is examined in Section 2 of the thesis ("Real-time monitoring of NH_4^+ with A-ISE sensors") and is the subject of **Paper I**. The final three questions are explored in Section 3 of the thesis ("Real-time forecasts of rainfall with NWP"), while **Paper II** specifically deals with question 3 and **Paper III** investigates question 4. Section 4 ("Discussion") provides a discussion of key results from the thesis, while Section 5 ("Conclusions") summarize the main findings of the PhD project. Finally, Section 6 ("Future research") points towards necessary and promising avenues of research.

2 Real-time monitoring of NH_4^+ with in-sewer sensors

2.1 A-ISE sensor use in wastewater management

Ammonium ion-selective electrodes (A-ISE) are one of the main options for continuous real-time monitoring of NH_4^+ concentrations in wastewater. They rely on ISE technology, which measures the electrical potential of a substance and relates it to the concentration of a target ion. ISEs contain a membrane that ideally only allows the specific target ion to affect the measuring electrode, and thus only the activity of the target ion in the otherwise complex mix of substances in wastewater is measured.

The main advantages of A-ISE sensors are that they can be installed in-situ (situated directly in the wastewater stream), and that they are significantly cheaper to purchase and operate than alternative ex-situ “analyzers” (Kaelin et al., 2008; Winkler et al., 2004). However, A-ISE is by many viewed as a less reliable technology, as several sources of uncertainty affect their measurements. Their membranes are never 100% exclusive to other ions that are similar to the target ion, and other ions can thus interfere with the measurements. Interference on NH_4^+ estimates in wastewater are mainly caused by K^+ and Na^+ (Cecconi et al., 2019; Winkler et al., 2004). ISE outputs are also known to drift over time with the degree of drifting depending on the specific sensor type and its usage (e.g. Papias et al., 2018). Their in-situ nature also means that the sensors are at high risk of clogging.

These issues can at least be partially mitigated through periodical cleaning of the sensor head, along with recalibration of the relationship between measured electrical potential and NH_4^+ concentrations against reference samples. However, improper or even over-zealous recalibration of a sensor might significantly deteriorate the final signal output, and using reference samples that do not adequately represent the current mix of wastewater can lead to poor recalibration outcomes (Cecconi et al., 2019).

The choice of standard operating procedure (SOP) for how an ISE sensor is maintained is clearly important. The data used in **Paper I** came from a measurement campaign conducted by the largest wastewater company in Denmark, which consulted with the sensor manufacturer on how to best care for it. The sensor was regularly cleaned with a wet cloth and recalibrated against a grab sample taken next to the sensor in the wastewater stream. The location of the

ISE sensor was close to optimal at a WRRF: behind the primary clarifier where many of the harsh constituents of raw wastewater have been removed (Winkler et al., 2004). This setup was state of the art data collection in terms of what can be expected from everyday use at a WRRF. The outcome was a yearlong time series riddled with sudden jumps due to improper sensor recalibrations, data that was difficult to use, and a utility company with serious distrust for this type of sensor. This highlights that collection of faulty data happens even under some of the best practical circumstances as wastewater is a harsh and difficult medium to sample from.

2.2 Handling faulty sensor data

There is a rich literature on fault detection and data quality control of water quality sensors (Leigh et al., 2019). These range from rule-based approaches, control charts, and mass balance models (e.g. Rieger et al., 2010; Thomann et al., 2002) to more advanced multivariate statistical approaches and principle component analysis (e.g. Alferes et al., 2013; Haimi et al., 2016).

One thing is detecting errors in data, another is what to do with a dataset that contains errors. Some often used ways of handling faulty sensor data are:

- *Discard errors*: A common approach is to simply discard faulty data points and only use the parts of a time series where data quality is considered reasonable. This avoids potential biases and wrong conclusions during data analysis and modelling, and abides by the philosophy that garbage inputs are going to lead to garbage outputs. It does, however, potentially leave holes in the dataset that might require filling and is also potentially a waste of resources.
- *Extract features*: It is possible to extract valuable features/signals from an otherwise unreliable dataset and make decision based on these. An example is Schneider et al.'s (2019) investigation of whether the detection of a trough in pH measurements, and an inflection point (where the second-order derivate is zero) in dissolved oxygen and nitrite concentrations could be used for monitoring small-scale, unstaffed treatment facilities. Here, it does not matter if a signal has an erroneous offset or has drifted as long as specific patterns of interest, e.g. a peak or a trough, can be detected.
- *Estimate and replace*: Gap filling techniques and so-called software sensors are trained to estimate what the actual value of missing or erroneous data points are. Such techniques can e.g. be based on interpo-

lation, historical data of the target variable, or through estimated relationships with other measured variables (De Mulder et al., 2018; Yang et al., 2020).

- *Quantify error and reconstruct*: Others try to quantify the size of the error in a given data point and subtract it to obtain an estimate of the actual value. An example is fitting a simple linear function to a drifting signal and subtracting the magnitude of the drift from erroneous raw A-ISE data (Papias et al., 2018).

2.3 Reconstruction of corrupted A-ISE datasets

Paper I developed a methodology that falls into the last category of how to handle erroneous data, and also tested whether the reconstructed data improved the training of a software sensor. Here, the reliability and usefulness of A-ISE data were increased by merging it with information from an additional data source. Volume-proportional composite samples were used as they are widely available at WRRFs. In Denmark they are required by law for reporting on treatment performance to the national regulators.

A-ISE data has the benefit of a continuous and high temporal resolution output (**Paper I** used two-minute frequency), but also has the downside of low accuracy. Composite samples have low temporal resolution (here 24-hour averages) but are deemed more accurate, hence their use for regulatory reporting. Figure 1 shows an overview of the steps that comprise the presented

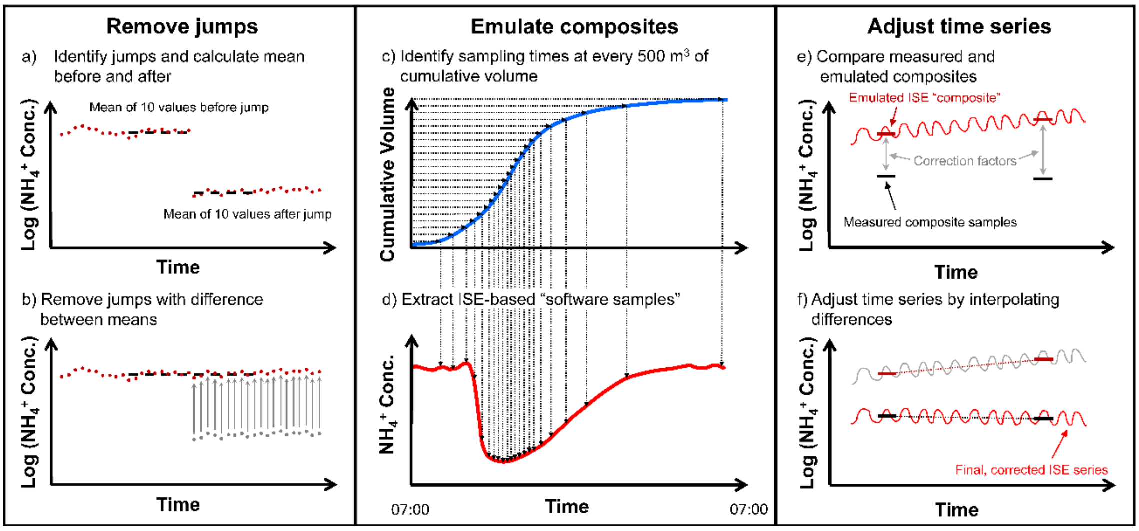


Figure 1: The sequence of steps involved in the developed correction methodology. Please note that the time scales of each column is different: a couple of hours (left), a 24-hour period (middle), two weeks (right) (Modified from **Paper I**).

methodology for merging the two datasets. First, all manual recalibration jumps are removed from the dataset (left column), and then the ISE data is sampled in a manner that emulates how the composite sample is constructed for the purpose of direct comparison (middle column). Finally, the ISE data is adjusted to fit the composite samples on the days that they were measured. See **Paper I** for a detailed description of the procedure.

Figure 2 shows two examples of how the composite samples have been used to correct the raw ISE signal. The left side shows a period with many poor sensor recalibration events that had led to a very erratic A-ISE signal, whereas the corrected signal did not contain these jumps. The corrected signal had dry weather concentrations in the range of 45-55 mg/L, which seems more physically realistic than the 30-80 mg/L variations seen in the raw data. The right side of Figure 2 shows a month where the raw signal slowly drifted away from the composite samples, while this effect had been removed by the correction methodology.

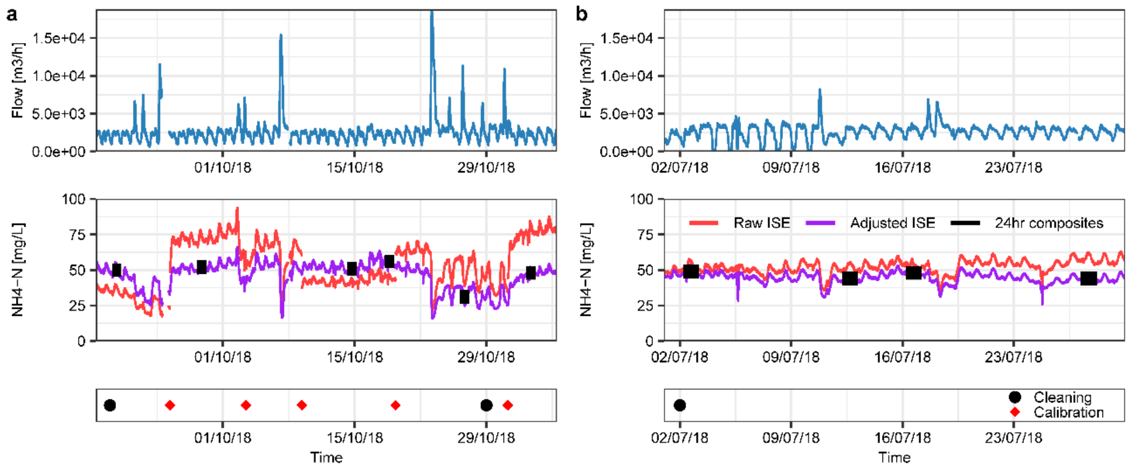


Figure 2: Example of two periods where the flow of water at the WRRF inlet is presented in top row. The middle row shows the raw and adjusted ISE signals, while the bottom row indicates sensor maintenance actions (**Paper I**).

One way to implement a software sensor is with a model trained to provide real-time estimates of a target variable given another measured variable as an input. **Paper I** examined the positive impacts that the data correction method had on training such a software sensor. Results showed that a software sensor trained on the reconstructed data had sharper parameter distributions and less uncertainty in the estimates of NH_4^+. Training the software sensor required eight weeks of A-ISE data to yield good median estimates, while 16 weeks of data were required for good predictive bounds (see **Paper I** for details).

3 Real-time forecasts of rainfall with NWP

3.1 Rainfall observations and radar nowcasting

3.1.1 Rainfall observations

Measurements of rainfall are typically performed at point locations with rain gauges and disdrometers, or through remote sensing with weather radars and satellites. For urban drainage applications, the most common data sources are rain gauges of various types as well as X- and C-band weather radars. While rainfall estimates from rain gauges certainly contain uncertainties (Ciach, 2003), they are often considered a highly reliable source of data and are often used as “ground truth”. Their main disadvantage is that they are point estimates of rainfall, and a dense network of gauges are required to adequate sample rainfall events with high variability (Peleg et al., 2018; Villarini et al., 2008). Weather radars scan the atmosphere by emitting microwave pulses and recording the backscattered reflectivities caused by hydrometeors such as rainfall. The reflectivities can be processed and converted into gridded spatial data. Rainfall intensities are estimated from reflectivities through empirical equations such as the Marshall-Palmer relation. The spatial dimension of weather radars is a large strength but the calculated rainfall intensities are highly sensitive to the raindrop size distribution. Weather radar data is therefore often merged with rain gauge observations for improved rainfall estimates (Goudenhoofdt and Delobbe, 2009; Ochoa-Rodriguez et al., 2019). Detailed reviews of weather radar use for urban drainage purposes can be found in Einfalt et al. (2004) and Thorndahl et al. (2017).

This thesis used rainfall observations from a rain gauge network at an urban catchment in Copenhagen. The rain gauge data is used for evaluating forecast performance in **Paper II** and calibrating a hydrological model in **Paper III.** Observations from Danish Meteorological Institute’s (DMI) national C-band radar network is used for visual classification of rain events in **Paper II**.

3.1.2 Radar nowcasting

A common method for generating short-term forecasts of rainfall is through so-called “radar nowcasting”, which rely on extrapolation of the observed spatial data provided by weather radars. There are many variants to how this is done. Generally though, consecutive radar scans are compared to each other and a vector field, which shows the trends of rainfall movement, is calculated e.g. through cross-correlation. Observations from the most recent radar scan are then advected along the vector field to provide estimates of future rainfall. More advanced nowcasting techniques account for factors such growth and decay of rain cells, and quantify uncertainties through stochastic perturbations (Bowler et al., 2006; Pulkkinen et al., 2019). Radar nowcasts are generally thought to provide skillful predictions at forecast horizons of 30 minutes to two hours ahead (Thorndahl et al., 2017). The main alternative to radar nowcasts when it comes to forecasting rainfall is numerical weather prediction (NWP), which is the main topic of this chapter.

Paper II of this thesis uses a simple radar nowcasting product as a benchmark for NWP-based rainfall forecasts. The radar nowcast is produced by DMI based on data from the national C-band radar network (see **Paper II** for details).

3.2 Numerical weather predictions

3.2.1 What are numerical weather predictions?

NWP are large-scale, physics-based simulations of atmospheric processes that attempt to predict the future state of the weather. For these simulations, the atmosphere is discretized into a three-dimensional grid where variables such as air pressure, density, temperature and winds are computed for each grid box through fundamental physical principles such as the laws of thermodynamics and the Navier-Stokes equations (Bauer et al., 2015).

NWP models require enormous computing power, which is reflected by national and international weather services having some of the largest supercomputer infrastructure in the world. However, NWP simulations are still limited by the available computing capabilities, despite the fact that computers have developed tremendously over the past decades. Various aspects of a NWP setup therefore trade-off against each other. Some of the most important aspects are spatial extent of the covered area, spatial resolution of grid boxes, size of integration time steps, forecast horizon, frequency with which new forecasts are made, and the number of members in ensemble prediction systems (EPS). Meteorological modelers balance these trade-offs differently depending on the purpose of a given forecasting system, which results in a range of different NWP products.

3.2.2 Different types of NWP

In terms of spatial extent, this gives rise to two types of models: global NWP models that simulate weather conditions for the entire planet at coarse spatial and temporal resolutions, and local area models (LAM) that only simulate regional conditions but do so in high resolutions. LAMs require initial and boundary conditions from global NWP models, and LAMs are therefore also referred to as being “nested” within a global model.

In terms of forecast horizon, various NWP models are designed to have high forecast skill for a targeted time window. Generally, forecasting setups are differentiated by the forecasting horizon as follows:

- Nowcasting: 0 – 6 h
- Short-range forecasting: 6 – 48 h
- Medium-range forecasting: 2 – 15 days
- Long-range/seasonal forecasting: > 15 days

Global NWP models often run once or a few times per day at operational weather services (seasonal forecasts may only be run once a month), while LAM NWP products where new forecasts are made frequently (often down to once an hour) are denoted “rapid updating cycles”, “rapid refresh”, etc.

Later in this thesis, Section 3.4 is going to review how NWP has been used in the urban drainage literature. To provide an idea of how some important properties vary between operational NWP setups, Table 1 shows a summary of these properties for the NWP products that have been used in the reviewed studies. The table shows that operational NWP setups range from global coarse-resolution models (70x100 km^2 grid boxes, 3-hour time steps, 10-day horizon, once per day updating) to very high-resolution LAMs (1x1 km^2 grid boxes, 10-minute time steps, 6-hour horizons, once per hour updating). For the studies that have used NWP ensembles, the number of ensemble members range from 10 to 26.

Table 1: Selected properties for NWP products used in urban drainage case studies.

Property	**Minimum**	**Median**	**Maximum**
Spatial resolution	1x1 km^2	3.3x3.3 km^2	70x100 km^2
Temporal resolution	10 min	1 h	3 h
Forecast horizon	6 h	31.5 h	10 days
Forecast frequency	1 h	6 h	24 h
Ensemble members	10	22	26

3.2.3 How NWP models conceptualize the world

The spatial and temporal resolutions of NWP models are central to how the physical processes of the atmosphere are conceptualized. Important physical processes that exist at scales smaller than a NWP model's grid discretization are accounted for by conceptual equations that act as source or sink terms in each grid box. This conceptualization is called "parameterization" in the meteorological modelling community. NWP models therefore either explicitly simulate or parameterize different atmospheric processes depending on a NWP model's resolutions. An important process such as deep convection, which frequently cause small-scale, high-intensity cloudburst events that might lead to urban pluvial flooding, exist on scales of 500 m to 10 km.

Scales and resolutions do not just constrain what a NWP model is able to resolve and simulate, they also have a large influence on how observations of rainfall appear. The variability of measured rainfall rates within a single rain event can be large even at small spatial scales, and it can have a large impact on urban hydrological modeling (see Cristiano et al. (2017) for a review). As an example of small-scale variations, Peleg et al. (2018) studied how average measured rainfall intensities within a 1x1 km C-band weather radar grid box can vary at any given point on the ground. They found that an extreme event with a measured grid box average of 150 mm/h could be observed as anywhere between 130 and 195 mm/h at point-scale (which is what a rain gauge on the ground would measure).

Figure 3 shows examples of what rainfall fields from a NWP model (3.3x3.3 km) look like compared to observed fields from DMI's national C-band weather radar network (500x500 m). These are the data products used in **Paper II**. The top row of Figure 3 shows a stratiform rain event with the center of a cyclonic low-pressure system at the top of the images, as well as several small showers in the southern half of Denmark. The bottom row shows a small convective rain event with intense rainfall near Copenhagen. The examples highlight that the NWP predicts the general weather patterns well, as both the center of the low-pressure system the showers are present in the top row example, and the convective cell is predicted near Copenhagen in the bottom row. However, the exact location of smaller rain cells are predicted less well. Small cells also gets smoothed by both the coarser spatial resolution of the NWP compared to the radar observations, and the fact that a NWP model requires approximately five adjacent grid boxes to resolve these phenomena (Golding, 2009). A small rainfall cell that in reality takes up the space of single grid box will therefore appear to take up a larger spatial extent in the NWP output.

The land-ocean discretization in the NWP model is also visible in Figure 3. The background map of Denmark for the radar observations is made with vector graphics that can show the coastline of Denmark in very high resolution (gray areas are land mass, white represent ocean). The background map for the NWP field is constructed from binary 0-1 values of how the weather model “sees” the shape of Denmark’s land mass in its roughly 3x3 km grid. It is in this resolution processes that describe ocean-air or land-air interactions (such as evapotranspiration) can be represented in NWP models.

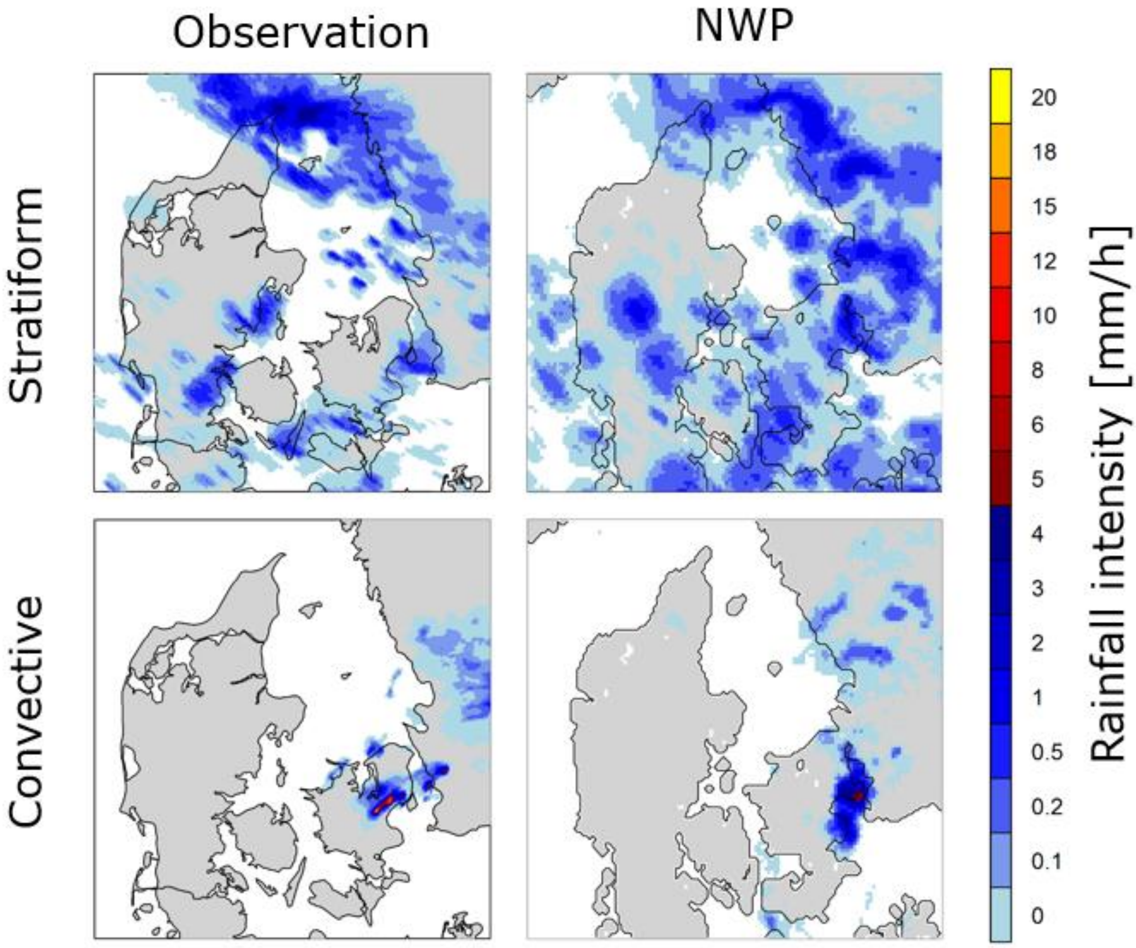


Figure 3: Example of a stratiform rainfall event with many adjacent, small showers (top row) and a small convective rainfall event (bottom row). The observed rainfall field for 10-minute average precipitation intensities from C-band weather radars (left) and the forecasted NWP field (right) over Denmark. The shown NWP and radar data products are the ones used in **Paper II**.

3.2.4 Ensemble forecasts

Despite the large progress that has been made within the field of NWP over the past decades, the atmosphere remains a highly unpredictable and chaotic system. Even small deviations in the initial conditions of a NWP model can lead to large differences in the forecasted values as the forecast horizon grows. This highlights that a single, deterministic forecast cannot describe the large uncertainties related to NWP. Ensemble prediction systems (EPS) have therefore been developed to address this and have become the meteorological standard for quantifying uncertainty in NWP. An EPS is essentially built from multiple Monte Carlo simulations of a NWP model where the initial conditions and/or model components are perturbed. The schemes that obtain ensembles by perturbation to the model are differentiated as (Du et al., 2018):

- "Multi-physics": Different parameterization schemes are used for selected processes within the same core NWP model. This accounts for uncertainty related to the choice of process conceptualization.
- "Stochastic physics": Stochastic perturbations are made to one or more components of a single NWP model. The perturbations can e.g. be made to parameter values, model states, and as random additive or multiplicative noise in specific equations.
- "Multi-model": Predictions from multiple different NWP models are collected into an ensemble. This approach accounts for both the choice of parameterization schemes like the multi-physics setups, and for uncertainty related to the choice of resolutions and the numerical integration schemes (if these are different between the various models that comprise the multi-model ensemble).

Figure 4 shows an example of an ensemble rainfall forecast over Denmark, which comes from a 25-member LAM EPS based on the DMI-HIRLAM-S05 model at approximately 5x5 km^2 resolution (Feddersen, 2009). This is the NWP product used in **Paper III**, and the ensemble is constructed as a mix of perturbations to initial conditions, multi-physics and stochastic physics. The figure shows a forecast issued on August 31, 2015, that predicted rainfall intensities 14 hours ahead. Most ensemble members agreed that a rainfall system would pass over western Denmark, but there were large variation in total rainfall depth and which regions of the country that would be most affected.

It is important to keep in mind that ensemble forecasts are developed such that the various members have diverged from each other in a way that shows good spread within a specific time window. The NWP ensemble shown in

Figure 4 is designed for short-range purposes, and its members are supposed to diverge from each other quickly. On the other hand, members of an EPS from a global medium-range NWP model, which targets >2-day forecast horizons, will not necessarily have diverged from each other in an adequate manner at horizons shorter than two days. The ensemble spread will therefore likely underestimate forecast uncertainty within the first two days.

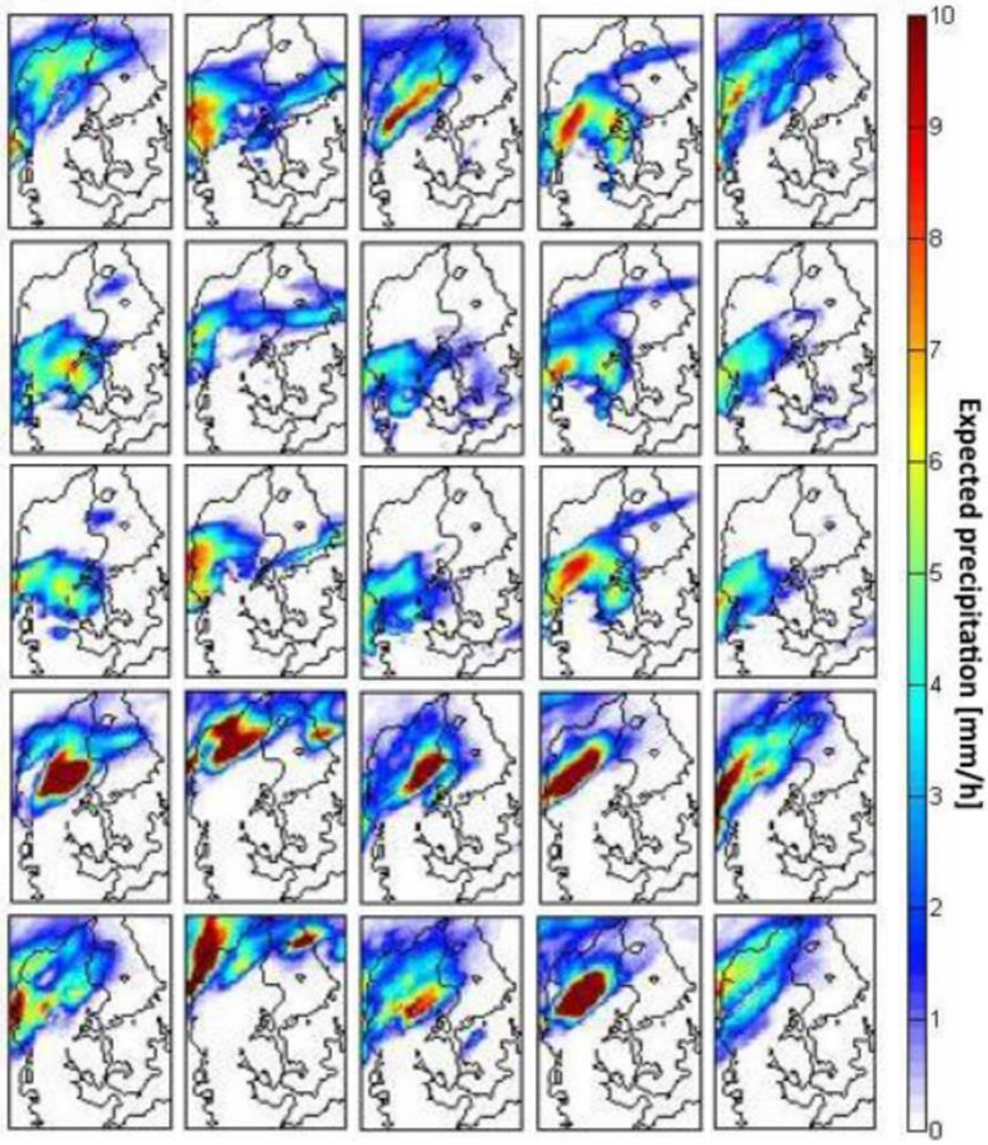


Figure 4: An example of a rainfall forecast from the DMI-HIRLAM-S05 25-member ensemble forecast for a rain event over Denmark. The snapshot shown in the figure is of predicted values 14 hours ahead given as accumulated mm over a one-hour time step. This data product is used in **Paper III** (modified from Courdent (2017)).

3.3 Forecast evaluation

Paper II and **III** evaluate forecasts with a range of metrics that are designed to highlight different performance aspects. The metrics apply to either deterministic point forecasts or ensemble forecasts, and some metrics are designed for categorical predictions. To understand the results presented in this thesis document, the most relevant metrics are those that pertain to categorical predictions.

Categorical metrics in their simplest form are binary yes/no evaluations of whether a forecast and an observation agree at a specific time. In **Paper II** these metrics were used to evaluate if a forecast could correctly predict when the observed rainfall had exceeded pre-specified thresholds of interest. In **Paper III** they were used to assess two flow thresholds: a low threshold signifying that stormwater was present in a combined sewer system, and a high threshold signifying CSO occurrence.

Table 2 shows a contingency table, which form the basis of many categorical evaluation metrics. If an observation or a forecasted value exceeds a threshold then they are counted as being "positive", and if the threshold is not exceeded they are considered "negative". There are four possible outcomes to a binary prediction. A true positive (TP) where both the observation and forecast exceed the threshold; a true negative (TN) where both do no exceed the threshold; a false positive (FP) where the forecast exceeds, but the observation does not; and a false negative (FN) where the forecast does not exceed, but the observation does. A contingency table is constructed by evaluating all issued forecasts and sorting them into one of these four outcomes. Performance metrics can then be constructed by counting the number of forecast outcomes in each category.

Table 2: A contingency table where a forecast outcome is sorted into one of four possible status indicators depending on whether it was a correct prediction (green fields), or a wrong prediction (red fields).

	Observation Positive	Observation Negative
Forecast Positive	∑ True Positives (TP)	∑ False Positive (FP)
Forecast Negative	∑ False Negative (FN)	∑ True Negative (TN)

The True Positives Rate (TPR) is a metric that describes the fraction of positive observations that were correctly predicted.

$$TPR = \frac{TP}{TP + FN}$$

The False Positives Rate (FPR) describes the fraction of the negative observations that were wrongly predicted as positive.

$$FPR = \frac{FP}{FP + TN}$$

The Positive Predictive Value (PPV) describes the fraction of positive forecasts that turned out to come true as positive observations.

$$PPV = \frac{TP}{TP + FP}$$

While TPR measures how reliable a forecast is at detecting the cases of interest, e.g. the percentage of CSO events that were correctly predicted up front. FPR and PPV are different ways of describing how prone a forecast is to making false alarms. Perfect TPR and PPV scores have values of one and the worst possible score is zero, while the opposite is true for FPR scores.

The three scores are interesting in their own right but can also be combined for graphical forecast assessments of probabilistic and ensemble forecasts. Relative operating characteristics (ROC) diagrams are constructed by plotting FPR and TPR on the first and second axes, while a Precision-Recall (PR) diagram is made by plotting TPR and PPV on the axes (Davis and Goadrich, 2006). An ensemble forecast will be evaluated as a curve in the two diagrams, which is constructed by plotting a series of points and drawing a line through them. The points are estimated by constructing as many contingency tables as there are members in the ensemble. The first contingency table is calculated by counting the outcomes of defining a forecast as positive if just a single member in the ensemble predicts a “positive” value. For the second table two ensemble members have to agree before a forecast is counted as positive, the third table requires three ensemble members to agree before it is positive, and so on.

Figure 5 shows a fictive example of how the performance of two competing forecast systems, F1 and F2, appear differently in the two diagrams. Good performance means that the ROC curve is pushed into the top left corner, while good performance means a PR curve pushed into the top right corner.

ROC analysis is commonly used in hydrometeorological assessments, but PR analysis is not. **Paper III** argues that this is a mistake as PR diagrams are more relevant for imbalanced datasets, which is often the case in urban drainage oriented performance evaluation. Here, an imbalanced dataset is understood as a dataset with few positive and many negative observations, which is the case for rare events such as flooding and CSO.

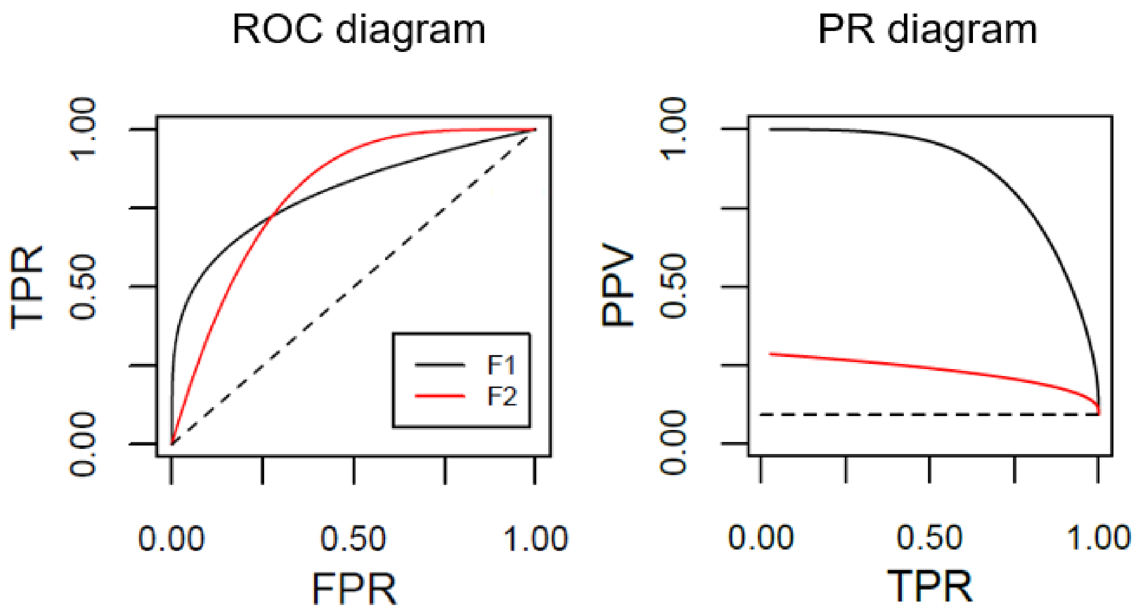


Figure 5: Illustration of ROC (left) and PR (right) diagrams. A perfect curve is pushed into the top left corner of the ROC diagram and the top right corner of the PR diagram. The curves correspond to two fictive forecast systems (F1 and F2) and show how their performance appear differently between the two diagrams. The dotted line depicts how a random prediction would perform (modified from **Paper III**).

3.4 NWP applications in urban drainage management

While rainfall forecasting based on NWP as a technology has been around for decades, the use of it for urban drainage and wastewater purposes is rather new. The first publications in the scientific literature started showing up in the early 2010's, and the collective literature on the subject is still sparse. This thesis has reviewed all ISI journal publications that employ NWP as an input for an urban drainage purpose, be it forecasting of in-sewer variables such as flow and water level, control of important actuators, warnings of urban flood inundation, etc. The total literature amounts to 20 publications, which can be categorized into four main topics:

1. Generic rain and flow forecasting
2. Urban pluvial flood forecasting
3. Real-time control
4. Post-processing

A summary of the state of the art is given below for each category as well as an explanation of how the results of **Paper II** and **III** fit into the various categories.

3.4.1 Generic rain and flow forecasting

A handful of urban drainage-related studies have investigated the use of NWP for predicting various variables in sewer systems, such as flows and water levels, and compared it against radar nowcasts as benchmarks. The scope and results of these are reviewed in this section.

Most of the published literature actually test forecast products that merge radar nowcasts with NWP, rather than evaluating the usefulness of NWP by itself. The first publications in the literature were two connected studies that examined the same forecast product where the STEPS algorithm (Bowler et al., 2006) was used to merge a deterministic NWP with a radar nowcast (Liguori et al., 2012; Schellart et al., 2014). The merged product was used to force a detailed hydrodynamic model for generating flow predictions in a small urban catchment. Liguori et al. (2012) found it difficult to produce flow predictions of high quality in general, while Schellart et al. (2014) concluded that the inclusion of the NWP improved the predictions for lead times longer than 1 hour and 45 minutes compared to the raw radar nowcast. In general, the NWP showed poor forecast accuracy at the small urban scales (Liguori et al., 2012), and the authors conclude that outputs with higher spatial resolution would be needed to forecast phenomena such as CSO and pluvial flooding (Schellart et al., 2014). Both studies struggle with obtaining strong conclusions as they relied on examining just three and five events, respectively. Another study also tested a product with merged radar nowcasts and NWP, where the forecasts consisted of radar nowcasts for the first 0-2 hours, a mix of radar and NWP for 2-4, and only NWP beyond 4 hours (Jasper-Tönnies et al., 2018). This was done after first seeing that the radar-only ensembles performed best for the first two hours of forecast horizon, while an NWP ensemble product were better beyond two hours for convective rainfall events. This merged ensemble product clearly outperformed the TPR scores of a reference consisting of a deterministic NWP. Yoon (2019) tested another technique for merging five different rainfall forecasts, three radar nowcasts and two high-resolution NWPs, into a single rainfall field over Seoul. The products were combined in a multiple linear regression where the weights were estimated based on the errors of the previous forecasts for each product. However, the merged product did not provide a clear improvement over the best individual forecasts for neither rainfall estimates nor sewer water levels after having been routed through a hydrodynamic model. The conclusiveness was also here hampered by an evaluation based on mere three rainfall events.

Urban runoff forecasts driven by pure NWP outputs have also been used for predicting rainfall depths at small urban scales and the inlet flow at a WRRF (Thorndahl et al., 2013). An evaluation of six rain events showed that radar-based forecast mostly outperformed their NWP-based equivalents for forecast horizons up to two hours. They also saw that the NWP performance actually improved for lead times of 6-12 hours compared to shorter forecast of 1-2 hours, which was likely because NWP models in general struggle with obtaining good initial conditions. Despite being a small study, their results were promising for WWTP inlet forecasting up to 24 hours ahead. Another study has shown that NWP products can have considerable skill even at small scales if the traditional flow forecasting problem is reframed from a question of predicting the exact amount of m^3/s to simply distinguishing between so-called “high” and “low” flow domains (Courdent et al., 2018).

In the context of the studies mentioned above, **Paper II** addressed several of the highlighted gaps and shortcomings. The results of **Paper II** were based on an archive of forecasts and observations containing more than 100 rain events, and its findings was thus much stronger than the less conclusive studies. Rainfall forecasts from a deterministic NWP were benchmarked against a standard radar nowcasting methodology. The examined NWP product is also of special interest to the urban drainage community as it assimilates radar observations during a warm-up phase for improved initial conditions, thus mitigating some of the issues related to very short-horizon predictions that e.g. Thorndahl et al. (2013) observed.

The forecasts were evaluated against ground observations from rain gauges at an urban catchment in Copenhagen, Denmark. This showed that both NWP and radar nowcasts had poor correlations with observed 10-minute rainfall intensities, and suggests that even state of the art rainfall forecasts are difficult to use directly at the scale of small urban catchments. The poor to mediocre performance from using NWP as direct input to hydrodynamic models (Liguori et al., 2012; Schellart et al., 2014; Yoon, 2019) can likely be explained by this.

Paper II also followed up on Courdent et al.'s (2018) finding that NWP can provide valuable skill in terms of discriminating between discrete flow domains (e.g. high vs low) rather than exact rainfall intensities. The local water utility companies in Copenhagen were asked to help delineate the rainfall domains that are most relevant for their operations, which resulted in four categories: (1) insignificant amounts of rain, (2) small rain events that likely do not cause issues, (3) medium-sized events that lead to bypass of wastewater at the WRRF, and (4) large events that can lead to CSO and surface flooding. Both NWP and radar nowcasts generally were good at predicting the cases without any rainfall, suggesting that they are well suited for determining if rainfall is going to occur or not.

Some of the studies mentioned above speculate about how the type of weather phenomenon that a rain event is a part of affects the predictive performance of NWP and radar nowcasts (Liguori et al., 2012; Thorndahl et al., 2013). **Paper II**'s large forecast archive allowed for a quantitative analysis of this based on a visual classification of 116 rain events in terms of four properties:

1. Evolution: How dynamically an event develops over time. Convective thunderstorms that arise and disappear quickly exemplifies high evolution.
2. Spread: The degree to which an event consists of small and scattered rain cells. In Denmark, small westerly showers are a common example of high spread, while a large, uniform frontal system has low spread.
3. Rotation: The amount of rotation in the incoming weather system. An example of high rotation are when the center of a low-pressure system moves directly across the case area.
4. Speed: The horizontal speed of the rain event. Large-scale frontal systems often do not move at very high speed, while smaller cells can be both fast-paced and slow dependent on the general weather conditions.

The results showed that a high degree of evolution and spread in a rain event had a negative impact on predictive performance of both NWP and radar nowcast (see Figure 6). Especially the radar nowcasts were severely impacted by this, as their simple advective extrapolation approach could not simulate these phenomena. The NWP with its simulation of atmospheric physics was better suited for these types of events. Rotation and speed of the weather systems did not have a clear effect on forecast performance. In general, the NWP was able to retain much of its skill throughout the 10-hour forecast horizon, while the performance of radar nowcasts rapidly declined as the horizon increased. The fact that the NWP can retain its performance for longer horizons makes it attractive for urban drainage problems that require long lead times such as control of storage volumes in sewer systems.

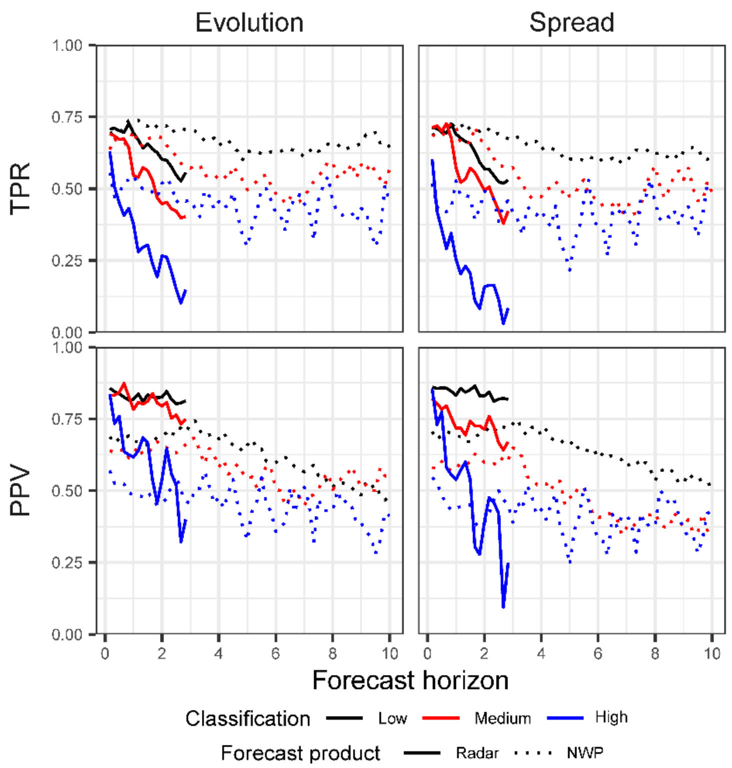


Figure 6: Average forecast performance over 116 events in terms of True Positives Rate (top row) and Positive Predictive Value (bottom row) for a rainfall threshold of 0 mm/h. Performance is shown as a function of the forecast horizon (x-axis) for two selected properties of the observed rainfall event (columns) based on a low-medium-high definition for each property (colors) (Modified from **Paper II)**.

3.4.2 Urban pluvial flood forecasting

There are two examples of published NWP-based pluvial flood forecasting systems in full operation, and they have taken different approaches on how to use information from NWP. Brendel et al. (2020) created a forecast setup that simply uses NWP rainfall fields directly as input to a hydrodynamic SWMM model. They tested their system on two events: a non-flood inducing event, which was correctly predicted as non-problematic, and a flood event, which the forecasts partially captured but which the local authorities likely would not have deemed severe enough to issue an official, public warning. René et al. (2018) attempted to use rainfall outputs from a global, deterministic NWP model with coarse spatial resolutions of 50x50 km^2 for a small urban case study (area of 0.45 km^2). They did not use the forecasted rainfall directly. Instead, their flood forecasts were based on 2D overland flow simulations using historical rainfall observations from previous floods as input. The measurement series from the historical events were scaled so their total rainfall depth equaled the depth of the incoming rain event predicted by the NWP model. René et al. (2018) tested their system on three flood events where it correctly issued warnings 12 hours ahead. However, some of their results suggested that false alarms might be an issue for their system, but it is hard to judge whether this is true without operational tests over a longer time period.

In lack of observational data, Yoon (2019) and Thorndahl et al. (2016) performed simulation studies where rainfall observations were used to force detailed 1D-2D flood models for one and two extreme rainfall events, respectively. The outputs of the flood simulations were used as pseudo-observations, which flood forecasts could be compared to. Rainfall forecasts from a deterministic NWP (Thorndahl et al., 2016) and a merged radar nowcast-NWP product (Yoon, 2019) were then used as inputs to the flood models. For Thorndahl et al. (2016) the results were significant underestimation of both rainfall intensities and flood extent, which the authors blame on the coarse spatiotemporal resolutions of the forecasts. The merged product also led to underestimation of flood extent, but less so than the raw NWP (Yoon, 2019). Other studies have tested NWP ensemble products with spatial resolutions less than 3 km, and thus in the higher-resolution end of the collective literature, as input to flood models (Jasper-Tönnies et al., 2018; Olsson et al., 2017). Jasper-Tönnies et al. (2018) found that the use of a NWP ensemble improved the number of correct flood warnings compared to a deterministic NWP over a three months period. However, direct use of a NWP ensemble in a conceptual hydrological model was not able to trigger a warning for a severe flood event in Malmö, Sweden (Olsson et al., 2017).

As explained above several studies have obtained poor to mediocre results through direct use of NWP in floods models, i.e. using the predicted rainfall intensities at the exact location of the catchment as input to the models (Brendel et al., 2020; Olsson et al., 2017; Thorndahl et al., 2016). This is not necessarily because the NWP products were unable to predict that intense rainfall was imminent, but due to the spatial uncertainty of the exact location of where the rain cells are going to hit. By including forecasted rainfall values from the immediate surroundings of their catchments, some authors have found small improvements to their otherwise failing forecasts (Thorndahl et al., 2016), while others obtained greater benefits (Jasper-Tönnies et al., 2018; Olsson et al., 2017). Olsson et al. (2017) even went so far as to make the spatial uncertainty a key component in their proposed decision framework, as they saw that several of their NWP ensemble members were able to predict intense rainfall but misplaced it in different directions. Consequently, they developed a visualization tool for improved flood risk assessment in a three-dimensional plot by showing median, minimum and maximum predicted accumulated rainfall depths as a function of distance from the urban catchment and the forecast horizon.

Yang et al. (2016) produced urban flood inundation warnings without a flood model by simply forecasting whether total rainfall depth would exceed predefined critical thresholds. Their system was tested on seven events, which showed that using the NWP forecasts for triggering warnings resulted in mediocre performance.

3.4.3 Real-time control

One of the main prospects of using NWP for urban hydrology purposes is for predictive RTC of drainage and wastewater systems. Despite this, the literature contains few case studies with RTC based on NWP and their results have been mixed.

Gaborit et al. (2013) investigated various RTC setups for the outlet gate of a stormwater pond receiving water from a small urban catchment. The aim of the study was to navigate a trade-off between increased settling of particulate matter in the pond while simultaneously avoiding overflows that could cause downstream flooding. The NWP was here used to predict if incoming rainfall would exceed the available capacity in the pond and if so, the excess volume would preemptively be discharged from the pond. However, the study showed little to no gain by using NWP compared to the reactive measurement-based setup. They suggest that their lack of good results were due to a poor choice of case study, and a follow-up study by the same authors therefore artificially modified the case to better investigate the effects of NWP-based control (Gaborit et al., 2016). They tested three different NWP products, but this study also showed virtually no benefit from using any kind of NWP.

Courdent et al. (2015) provided a simulated “proof-of-concept” case study with control of storage basins in an urban catchment based on the same LAM NWP EPS that is used in **Paper III**. The NWP product was used as an input to the MPC algorithm “DORA“ (Vezzaro and Grum, 2014), which is designed to minimize the risk of combined sewer overflows based on the relative costs of emissions from each overflow location in a catchment. The NWP outputs were used to predict which of three operational modes, each with different MPC objectives, that would be expected in the near future. It is worth noting that they only simulated a single CSO event, which nonetheless showed that the NWP-based MPC successfully diverted CSO occurrence from three expensive, upstream locations to a cheaper, downstream location. Their analysis showed that this was achieved due to the NWP’s long forecast horizons, which provided lead times long enough to pre-actively empty more stored water.

A couple of papers have investigated the use of NWP data for control of integrated urban drainage-wastewater systems (Courdent et al., 2017; Stentoft et al., 2020). The main idea is that WRRFs can benefit from optimal use of the storage capacity in the upstream sewer systems by temporarily retaining water in storage basins. The wastewater can then be released later when the WRRF is ready for it. Since WRRFs are vast consumers of electricity there are potential gains in energy savings for the facilities. In regions with large fluctuations in renewable energy production, such as the wind power-heavy Danish energy market, WRRFs can exploit dynamic energy prices and increase a region’s demand-side flexibility (Brok et al., 2020; Stentoft et al., 2020). It is, however, important that water stored in the sewer system does not limit the capacity to convey stormwater during rain events, and thus create a risk of surface flooding or CSO. Rainfall predictions are therefore necessary to ensure that any stored water can be removed before an incoming rain event arrives. Some energy markets rely on price prediction for the following day (e.g. NordPool, www.nordpoolgroup.com), and the required rainfall forecast horizons are therefore in the order of 1-2 days, which is suitable for short-range NWP.

Courdent et al. (2017) provide a simple methodology for determining whether predictive control based on NWP ensembles is suitable for a given case. The method uses the Relative Economic Value (REV), which depends on specifying the ratio between potential benefits from energy savings in relation to the potential costs of negative impacts from a missed forecast, e.g. CSO and flooding. The framework can then be used to determine the optimal decision-rules, such as how many ensemble members that need to agree that no rainfall is incoming before a switch to an energy optimization scheme is made. Stentoft et al. (2020) developed a MPC algorithm that optimized both electricity consumption and effluent quality of a WRRF. This was done by using the storage basin immediately upstream of the facility to control the inflow to the treatment processes. A rainfall forecast from a NWP determined whether the overall system control should switch between the developed MPC and a rule-based control during dry and wet weather, respectively. As the only real-life case study in the literature, they showed results from a full-scale test over seven days. However, it seems that no rainfall occurred during their seven days of full-scale operation, which makes it difficult to assess how the uncertainty in the rainfall forecasts will affect the performance of the algorithm over longer periods of time.

Paper II expands on the limited number of examples with NWP-based predictive control in the literature. Here, a NWP product was used to control a wet weather switch at a WRRF in Copenhagen. The WRRF treats wastewater during dry weather but also receives stormwater from a combined sewer system when it rains. The facility is able to increase its hydraulic capacity during wet weather through the technique of aerated tank settling, which protects against sludge escape from the secondary settlers by pumping some of the sludge into the aeration tanks (Sharma et al., 2013). The setup requires time to transition into a fully operational wet weather mode, which for this study was assumed to be around one hour. The examined NWP product was therefore used to trigger the switch from dry to wet weather mode if more than 1 mm of rain was predicted within a forecast horizon of 1.5 hours.

As shown in Figure 7, **Paper II** was able to show a clear benefit to using NWP-based rainfall forecasts compared to a reactive decision setup based only on rain gauge measurements (labelled “No forecast” in the figure).

Many more of the simulated events were in the acceptable categories of “good” and “early”, with the NWP-based scenario adequately predicting almost as many events as the “Perfect forecast”-scenario. However, the two-year analysis also showed that the NWP produced a large amount of false alarm switches. Whereas several of the other NWP-studies in the urban drainage literature have shown that radar nowcasts outperform NWP at horizons of 1-2 hours (Jasper-Tönnies et al., 2018; Thorndahl et al., 2013), **Paper II** found that a NWP product with improved initial conditions could compete with radar nowcasts in this kind of RTC setup (see Figure 7).

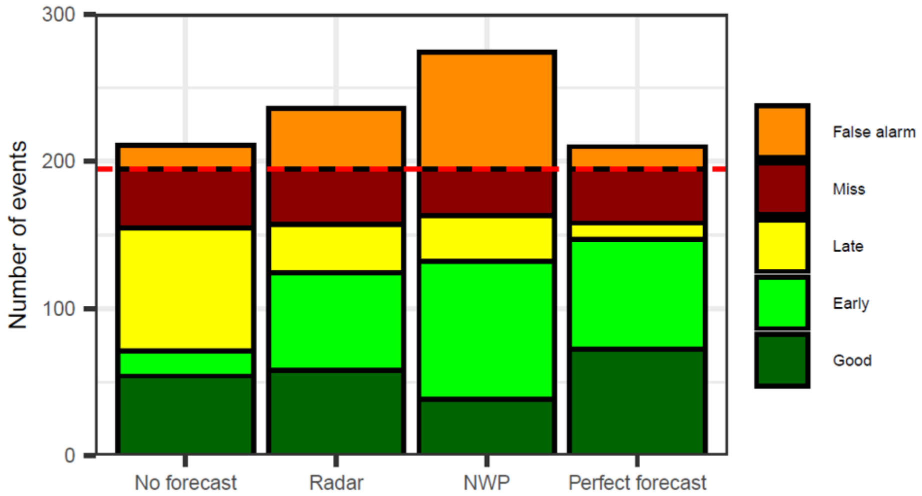


Figure 7: The status of how the wet weather switch performed for all rain events within a two-year period at the Damhusaaen WRRF. The dashed, red line indicates the total number of observed events (good + early + late + miss), while the predicted false alarm events are on top of these (**Paper II**).

3.4.4 Post-processing

NWP models often exhibit some degree of bias (as also seen in **Paper II**). NWP ensembles are also often underdispersive in the sense that they underestimate uncertainty and too many rainfall observations fall outside of the ensemble spread (Buizza et al., 2005), and ensembles scenario do not correspond to probabilities directly. These issues have motivated the development of so-called “post-processing” techniques, which aim at enhancing or correcting the raw NWP output.

Within the urban drainage literature, two types of post-processing techniques have been applied to NWP-based rainfall forecasts before further use: (1) statistical methods which rely on estimating a statistical model for correcting the forecast, and (2) in-expensive methods that expand the ensemble size to include more potential scenarios of future rainfall.

Two related papers have investigated statistical approaches where the first paper developed the post-processing method, while the second tested it in an operational setting (René et al., 2013, 2018). The method was able to produce probabilistic forecasts from a single deterministic NWP output, and did so by estimating conditional probability distributions from a comparison of previous forecasted values with their corresponding observations from a historical data archive (René et al., 2013). The distribution were estimated by fitting a bivariate Gaussian distribution, and was tested by René et al. (2018) for pluvial flood forecasting as described in Section 3.2.2. They found that the post-processing method improved TPR scores but led to a large increase in FPR. It is not clear that the method provided much value, as the raw and corrected forecasts produced almost identical warning decisions.

No other studies in the urban drainage literature have tested statistical post-processing methods, despite the fact that the meteorological literature contains numerous approaches such as regression-based methods (Messner et al., 2014), ensemble model outputs statistics (Scheuerer, 2014), and Bayesian model averaging (Raftery et al., 2005). This might be due to the issue that fitting statistical models requires a historical archive of forecasts and observations, which rarely has been available in the past. If any changes to an operational NWP model are made that affect how it predicts rainfall, then a new archive of forecasts will have to be produced by re-running historical events. Such changes can happen frequently as weather services continually tune their NWP setups for improvements (Vannitsem et al., 2020). The fact that extreme flood-inducing precipitation rarely occurs at any given location makes it even more difficult to properly estimate a statistical model for such events. It might therefore be more feasible to employ statistical post-processing for frequent everyday types of rainfall events where many observations and forecasts are available within relatively short time windows. This could for instance improve the RTC setups that rely on simple dry vs wet weather forecasts as mentioned in Section 3.2.3.

The remaining urban drainage studies that have used post-processing have all been neighborhood methods, which attempt to address issues of spatial uncertainty in the rainfall forecasts. The general approach of neighborhood methods starts by defining a large area surrounding the catchment (i.e. its "neighborhood"). Precipitation that is forecasted outside of the catchment but within the neighborhood area is considered as a potential rainfall scenario that might occur over the catchment in case there is spatial misplacement of a rain event. The method originated in the meteorological literature with Theis et al. (2005) as a way of generating an ensemble forecast from a deterministic NWP, while others since have used the method to create large super-ensembles from raw NWP ensemble products (Ben Bouallègue et al., 2013; Schwartz et al., 2010). An additional advantage of this method is that it is straightforward and computationally cheap to implement, and does not require a large historical forecast archive as some of the statistical methods. The urban drainage literature has used the neighborhood method to expand the number of ensemble members and thereby also the number of scenarios that can trigger a decision (Courdent et al., 2017, 2018). Others have used it for flow and CSO predictions (Courdent et al., 2018) as well as flood warnings (Jasper-Tönnies et al., 2018; Olsson et al., 2017) by evaluating whether high-intensity rainfall is predicted within a neighborhood. The size of the neighborhood is an engineering parameter to be tuned for the individual use case with the mentioned studies using maximum neighborhood sizes of 12.5-50 km.

An alternative inexpensive method is so-called lagged or time-lagged ensembles, which uses previous forecasts together with the newest one in an ensemble (**Paper III**). The justification for this is that the newest, most recently issued forecast is not necessarily better than the previous ones. This sometimes happen with NWP models due to poor initial conditions and spin-up effects caused by instabilities in the numerical model in the beginning of the forecast horizon. Like with neighborhood approaches, time-lagging is intuitive, originates from the meteorological community, and is computationally cheap and straightforward to implement. It also does not require a large forecast archive for implementation. Time-lagged forecast have been used in the meteorological community for creating ensembles from a series of single, deterministic NWP runs (Mittermaier, 2007) and for developing super-ensembles from a series of NWP ensemble runs (Ben Bouallègue et al., 2013).

Paper III distinguished between two ways of employing time-lagged ensembles. The first approach, “consistent signal”, required multiple time-lags to show exceedance of a threshold before an action was made. This approach is visualized in Figure 8 for an observed CSO event. For this event, the consecutive forecasts are quite consistent in predicting that CSO will occur, which should provide a forecaster with reassurance of what will happen in the near future. On the other hand, disagreement between consecutive NWP forecasts would highlight that there is large uncertainty about an incoming event (Pappenberger et al., 2011). The second approach, “cumulative super-ensemble”, aggregated the ensemble members of the individual forecasts into a common, larger ensemble with 50 to 125 members, which gave more scenarios of what might happen in the future.

Time	04.00	05.00	06.00	07.00	08.00	09.00	10.00	11.00	12.00	13.00	14.00	15.00	16.00	17.00	18.00	19.00	20.00	21.00	22.00	23.00	00.00	01.00	02.00	03.00
Obs																								
Lag4	0	0	0	0	0	0	0	0	0	1	9	22	24	24	21	4	1	0	0	0	0	0	0	0
Lag3	0	0	0	0	0	0	0	0	0	3	5	13	23	24	13	0	0	0	0	0	0	0	0	0
Lag2	0	0	0	0	0	0	0	0	2	5	12	18	20	16	4	0	0	0	0	0	0	0	0	0
Lag1	0	0	0	0	0	0	0	0	0	0	4	11	14	5	2	1	0	0	0	0	0	0	0	0
Lag0	0	0	0	0	0	0	0	0	0	0	5	13	19	19	11	4	0	0	0	0	0	0	0	0

Figure 8: Example of the information on forecast consistency that time-lagged NWP ensembles can provide to an end-user. Each row highlights the number of ensemble members that predict CSO occurrence in each of the five most recent forecasts, with the oldest forecast in the top row and the newest in the bottom row. The numbers in each row indicate the number of ensemble members in each forecast that predict CSO occurrence at a given point in time. The red bar in the “Obs” row show the hours where CSO was actually observed (**Paper III**).

The two methods were used to predict exceedance of two flow thresholds in a combined sewer catchment: 4000 m^3/h signifying the difference between dry and wet weather, and 9500 m^3/h signifying CSO occurrence. ROC and PR analysis showed that the cumulative super-ensemble performed best for both thresholds, while the consistent signal approach failed to improve on the raw, original ensemble.

Time-lags and neighborhoods address different sources of uncertainty in NWP models. As mentioned above, time-lags account for uncertainty in initial conditions and spin-up issues, while neighborhoods address spatial uncertainty. **Paper III** therefore compared the cumulative super-ensemble to the “maximum threat” neighborhood method (Courdent et al., 2018), which uses the highest intensities predicted within a specified neighborhood for each ensemble member. The two methods performed similarly for the 4000 m^3/h threshold (Figure 9a,b) but the cumulative super-ensemble covered more of the ROC and PR space, which means that an end-user has more points to operate from. For CSO predictions (Figure 9c,d), it is seen that the two curves cross each other in PR space, which suggests that the best method depends on the end user’s risk tolerance. If detecting many events (at the cost of more false alarms) is preferable, then the time-lagged method is preferable to the neighborhood method. The opposite is true if false alarm actions are expensive, and it is acceptable to have a somewhat lower detection rate to avoid these.

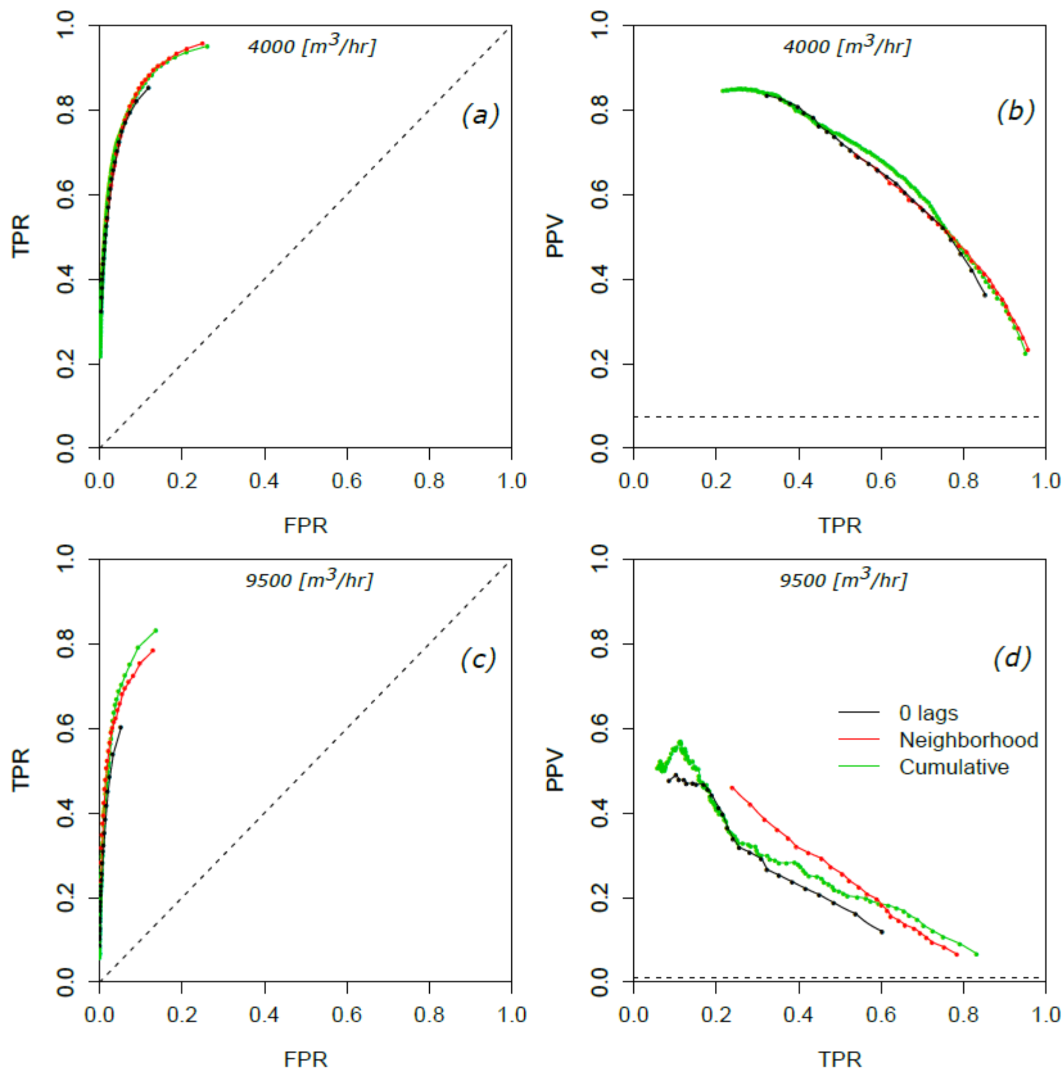


Figure 9: Comparison of the cumulative super-ensemble approach based on time-lagged forecasts and the maximum-threat neighborhood method. ROC (a+c) and PR (b+d) diagrams illustrate skill for flow thresholds of 4000 m3/h (a+b) and a 9500 m3/h (c+d) (**Paper III**).

4 Discussion

4.1 A-ISE sensors for monitoring in wastewater

4.1.1 How to operate A-ISE sensors

The raw A-ISE data collected for **Paper I** had serious flaws in terms of erratic recalibration jumps and drifting values. As we considered the setup to be state of the art in terms of what is practically achievable, there are serious concerns regarding the applicability of A-ISE sensors for wastewater monitoring. The WRRF operators that were in charge of managing the sensor during the study had little faith in the technology after seeing the raw data. It is clear that SOPs for A-ISE maintenance have to be updated as the currently recommended grab sample adjustments are inadequate for sensor recalibration. Cecconi et al. (2019) solved the issue with ex-situ recalibrations, i.e. grabbing a bucket of wastewater from the raw stream, putting the sensor in, waiting for it to adjust to the new medium, and taking a laboratory sample for comparison purposes. While this significantly improved their sensor readings, the operators we worked with consider the “bucket method” laborious and inefficient.

Paper I tested the usefulness of an A-ISE sensor at probably the most favorable practical conditions: behind the primary clarifiers. It is not clear how useful these sensors are in less optimal locations and for setups where high-quality reference samples for recalibration/merging are not available. Sewers and WRRF inlets may be too harsh an environment leading to issues with debris and fat coating (Winkler et al., 2004), while WRRF effluent and recipient surface waters may contain too low concentrations for ideal use (Papias et al., 2018; Winkler et al., 2004). It is also likely to be more difficult to establish high-quality reference samples outside of the WRRF, such as sewers and recipients.

Integrating the 24-h volume-proportional composite samples into the data setup solved many of the negative aspects with operating A-ISE sensors at the considered location. It did so without any additional operational costs as it exploited existing infrastructure. The new methodology thus provides a major boost to A-ISE use in wastewater monitoring.

4.1.2 Uncertainties in the data

While **Paper I** showed that merging the raw A-ISE data with 24-h volume-proportional composite samples provided a much improved quality of data, the methodology cannot account for all types of errors. Successful implementation at a WRRF requires a certain minimum standard of the raw data. In **Paper I**, the composite samples were regarded as a reliable source of reference data, which was motivated by the fact that they are used for regulatory compliance assessment. However, any type of sampling will be subject to uncertainty arising from the specific sampling technique, equipment failure, and laboratory analysis (Ort et al., 2010). While some of these uncertainties may be accounted for in a more advanced data merging algorithm, it will be highly important that the samples undergo rigorous quality control since they have a large influence on the final signal output.

Even though the developed methodology was able to remove some errors (jumps and drifting), the underlying raw A-ISE signal also have to provide meaningful raw estimates of NH_4^+ variability over the course of a day. The sensor will e.g. have to be protected against debris leading to major clogging issues. The sensor head will also still have to be cleaned regularly, e.g. once a week (Cecconi et al., 2020). The merging algorithm replaced the manual sensor recalibrations through what essential amounts to a so-called "one-point" calibration. Here, only the offset parameters in the relationship between electrical potential and NH_4^+ concentrations are updated. It is not clear to which degree the slope parameter is affected by sensor wear. Ohmura et al. (2019) showed that the slope parameter of ISE sensors for pH measurements did not change significantly over time, but it is an open question whether their results transfer to outside of their controlled settings and to A-ISE sensors. Field experiments suggest that the sensors may need to undergo a more comprehensive recalibration of both parameters about every three months (Cecconi et al., 2020).

4.1.3 Perspectives for A-ISE use

The reconstructed ISE data can be used for improved performance assessment, legal reporting to authorities, and modeling efforts such as influent generation (Langeveld et al., 2017; Martin and Vanrolleghem, 2014). It is also well-suited for training NH_4^+ software sensors and forecasting models (Newhart et al., 2020; Vezzaro et al., 2020). Finally, it can be used for feed-forward control allowing e.g. aeration processes to be determined by upstream NH_4^+ estimates rather than delayed information from effluent concentrations (Kaelin et al., 2008; Stentoft et al., 2019).

4.1.4 Expanding the methodology to other variables

The idea of merging various types of sensor data products with different appealing aspects should be highly transferable to other use cases. The Danish wastewater regulations require that composite samples are analyzed for other variables than NH_4^+ such as suspended solids, biological oxygen demand, nitrate, and phosphate. Low-cost online sensors of some of these variables could also be merged with the available composite samples. Nitrate would be a low hanging fruit since ISE sensors are available for it as well, often even combined with ammonium estimates in the AN-ISE sensor type.

Data merging setups are something that water utilities could, and perhaps should, plan for when they design their monitoring strategies. Other fields already do this with an example from urban hydrology being merging of rain gauge and weather radar data. The presented methodology improves the reliability of a low-cost sensor placed in the same location as a high-quality reference data source. However, it could also be possible to improve low-cost signals with reference data from another location by constructing a mathematical model that can describe how the dynamics of the two locations relate to each other. This may be promising for sensing in distributed sewer networks where many low-cost sensors can provide distributed (noisy) data of variables such as water level, flow, and various water quality parameters. This could be then updated with information from a few accurate sensors in key locations.

4.2 The future of NWP-based rainfall forecasting

4.2.1 An open future

In the past years, there has been a wave of opening of data products from national meteorological institutes that are now available free of charge. Many countries now provide free access to NWP products on multiple spatiotemporal scales and resolutions. These include but are not limited to the UK, France, the Netherlands, Germany, Sweden, Norway, Finland, USA, and Canada. It is also possible to obtain local observations and predictions from international centers such as EUMETNET and ECMWF. From 2019 onwards, DMI has also started to gradually release various observations and forecast products covering Denmark.

4.2.2 Next-generation NWP models: high-resolution, convection-permitting, and improved DA

Deep convection is an atmospheric process that happen at a small spatial scales (< 10km) and have previously been parameterized in NWP models. However, the new generation of NWP models, that are approaching 1 km spatial resolution, are better able to reproduce deep convection processes, and have become known as “convection-permitting” models. Clark et al. (2016) gives an assessment of what the new generation of convective-permitting models will do for the quality of NWP-based rainfall forecasting. For the near future, simple increases in computing power will continually improve NWP outputs by allowing more sophisticated and demanding DA schemes. Ensemble sizes can be increased for better uncertainty quantification, and forecasts can have longer horizons and larger covered domains. Increasing resolutions will allow more and more small-scale phenomena to be properly resolved with deep convection being an important one. These gradual developments have been termed “the quiet revolution” of NWP (Bauer et al., 2015), and it will continue to evolve in the coming years.

Convection is of special interest to urban hydrology as it can lead to high-intensity rainfall events that cause pluvial floods in cities. It is therefore excellent news that these types of events will appear ever more realistically in new NWP models. However, fully resolving all types of convective events in operational forecast products does not seem possible in the near future as that might require grid resolutions finer than 100 m (Bryan et al., 2003; Clark et al., 2016). While convective events now look more realistic in the NWP output, there will still be large spatial uncertainty associated with these events and ensemble sizes will remain too small to fully sample the range of possible outcomes. Post-processing methods such as neighborhoods and time-lags will therefore continue to be valuable, and will require continued application-oriented refinement (Clark et al., 2016).

RTC schemes for small urban catchments as well as those that do not have any significant storage facilities for retaining water require good rainfall predictions with short horizons. NWP products have struggled with good performance on such short horizons and radar-based nowcasts have therefore been preferred in the past. The NWP community will need to address the poor initial conditions mainly caused by lack of and insufficient assimilation of observations (Bauer et al., 2015). Several national weather services are developing various DA techniques that assimilate weather radar data into high-resolution NWP models (e.g. Ballard et al., 2016; Korsholm et al., 2015). **Paper II** used one of these NWPs with assimilation of observed radar reflectivities through latent heat nudging, which is one of the more commonly used methods (Gustafsson et al., 2018). Ongoing research will continue to investigate more sophisticated methods such as 4D-Var and adapting the EnKF for convective-scale NWP (Sun et al., 2014). These efforts will be critical for many urban applications.

4.2.3 Merging radar nowcasts and NWP

There is a large, ongoing research effort towards integrating radar nowcasts and short-range NWP into one, single, "seamless" forecast. The rationale for this is that radar nowcasts perform well on very short horizons between 30 minutes to 2 hours depending on the weather type, while NWP is better for longer horizons.

Some of the urban drainage studies that employ merged radar nowcast-NWP products have used rather simple merging techniques that rely on a weighted average of the rainfall fields (Jasper-Tönnies et al., 2018; Yoon, 2019). That might be a reasonable approach for large-scale, stratiform rainfall systems where radar nowcasts and NWP tend to largely agree. However, for convective events, which both of these studies focus on, averaging two fields with spatial disagreement of where rain cells are located can make it seem like there are twice as many rain cells in the merged forecast. Other studies (Liguori et al., 2012; Schellart et al., 2014) have used the somewhat more sophisticated STEPS algorithm (Bowler et al., 2006), which decomposes the two original forecasts into various spatial scales and accounts for various sources of uncertainty.

Other merging techniques that do not contain the potential pitfalls of simple averaging do exist in the meteorological literature. An example is an EnKF-based solution where a stochastic nowcasting technique starts off from a radar nowcast, which is gradually updated and smoothly transitions into looking like the NWP at longer horizons (Nerini et al., 2019).

4.3 Matching scales and resolutions of project aims, data, and models

One of the keys to successful implementation of a NWP-based forecasting or control system is proper recognition and alignment of the scales and resolutions of the project purpose, employed NWP product, hydrological model, and the decision framework. Of the surveyed urban drainage studies, those that fail to align these components are usually also those that struggle with obtaining good results. In this section, some common issues seen in the literature will be addressed.

4.3.1 Time scales of use cases and NWP forecast horizons

As mentioned earlier, NWP products are often designed with a given purpose in mind and to be most effective within a specific time window (nowcasting, short-range, medium-range, seasonal). It is therefore essential for successful use of NWP that urban drainage researchers and professionals employ the right kind of NWP product for the right purpose. Figure 10 compares the relevant time scales of urban storm- and wastewater management issues with the forecast horizons of selected rainfall forecasting products that are provided for Denmark by DMI.

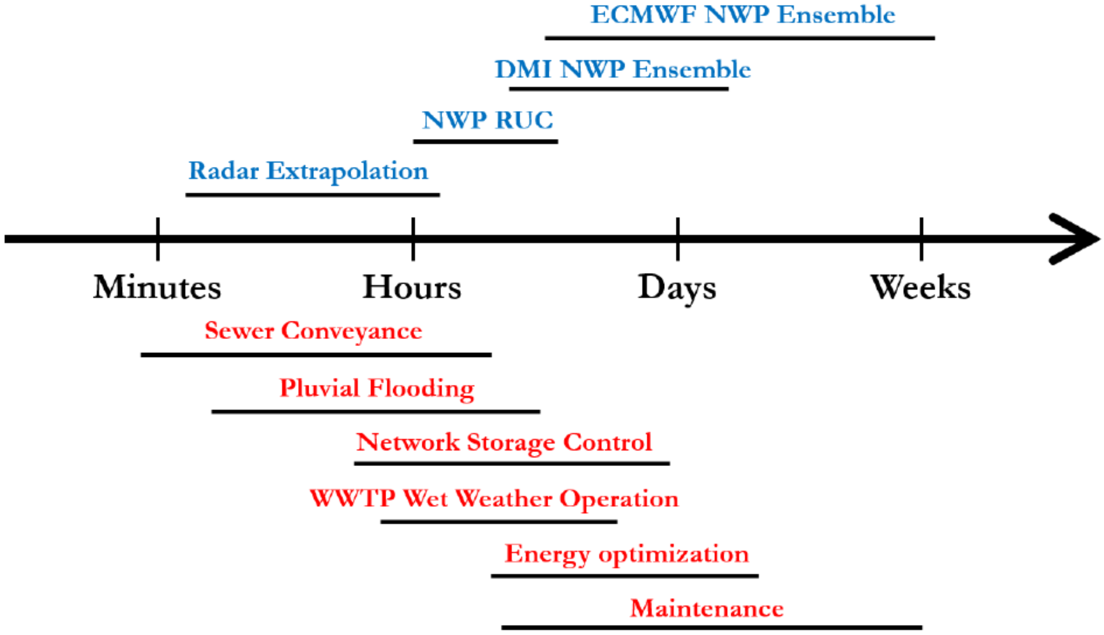


Figure 10: An overview of the forecast horizons of selected rainfall forecasting products that DMI provides for Denmark (blue), and the relevant time scales for a range of different aspects and objectives of urban storm- and wastewater management.

While there is little risk of someone using a short-range NWP product for a drainage purpose that exist on longer scales, since these obviously are insufficient, the literature does show examples of the opposite. As Figure 10 highlights, the use of a global, medium-range NWP, such as ECMWF's, with the goal of accurate in-sewer flow predictions or inundation depths a few hours ahead is equally out of proportion.

4.3.2 Spatial scales of urban catchments and skillful NWP outputs

The predictability of rainfall systems and thus the performance of NWP-based rainfall forecasting is highly scale-dependent (as seen in **Paper II** for the low performance of high-spread events). In general, larger precipitation systems tend to be more organized, have longer lifetimes, and are more predictable (Sun et al., 2014). Meteorological modelers have recognized that evaluating NWP rainfall at point locations can make the forecasts appear poor, while they actually contain large value and skill from a larger spatial perspective. This has encouraged a move towards spatially aware evaluation metrics such as the Fractions Skill Score, where skill is calculated at multiple scales (Ebert, 2009; Roberts and Lean, 2008), and neighborhood post-processing methods as described earlier. It is often the case that urban catchments have a size that fits within one or a few NWP grid boxes and through the wide lens of regional or global NWP forecast domains, these catchments are more or less points in space. Urban drainage users of NWP estimates must therefore take care before simply using the predicted values in the grid box(es) that their catchments fit within. Robust warning and control schemes will have to be aware of the spatial uncertainty in their rainfall inputs (Courdent et al., 2017, 2018; Jasper-Tönnies et al., 2018; Olsson et al., 2017).

4.3.3 The choice of hydrological models

Total runoff in urban areas is characterized by a high degree of surface runoff and very fast response times due to the many impervious surfaces and piped flow paths. Detailed 1D and 1D-2D hydrodynamic models are often used to simulate flows through the system and the urban drainage community have come to expect high-resolution rainfall products to force these models. General advice on rainfall data in the literature is to have spatial grid resolutions less than 1 km^2 and temporal resolutions less than 5 minutes (Ochoa-Rodriguez et al., 2015; Schilling, 1991).

Some studies use detailed, computationally expensive runoff models, which consequently only allow for simulation of a single to a few rainfall scenarios (Brendel et al., 2020; Liguori et al., 2012; Schellart et al., 2014; Thorndahl et al., 2016). These approaches rely on receiving highly accurate rainfall forecasts. Given the kilometer-scale spatial resolutions, the hourly temporal resolutions, and the large uncertainties related to most NWP data, such accuracy does not exist in practice. The direct applications of NWP in detailed pluvial flood models cannot be justified given that convective events are some of the most difficult to predict.

Other ways of using the information from NWP are clearly preferable, but using some sort of runoff model may still be valuable. It lets a forecaster account for catchment specifics that might be important for runoff production such as land use, soil moisture, system actuators, antecedent conditions in the drainage network, etc. In general, it makes intuitive sense that detailed rainfall products should feed detailed hydrological and hydrodynamic models in order to handle problems that require detailed information. Conversely, coarse and uncertain rainfall products should feed simplified runoff models for simplified decision-problems. Figure 11 visualizes this idea.

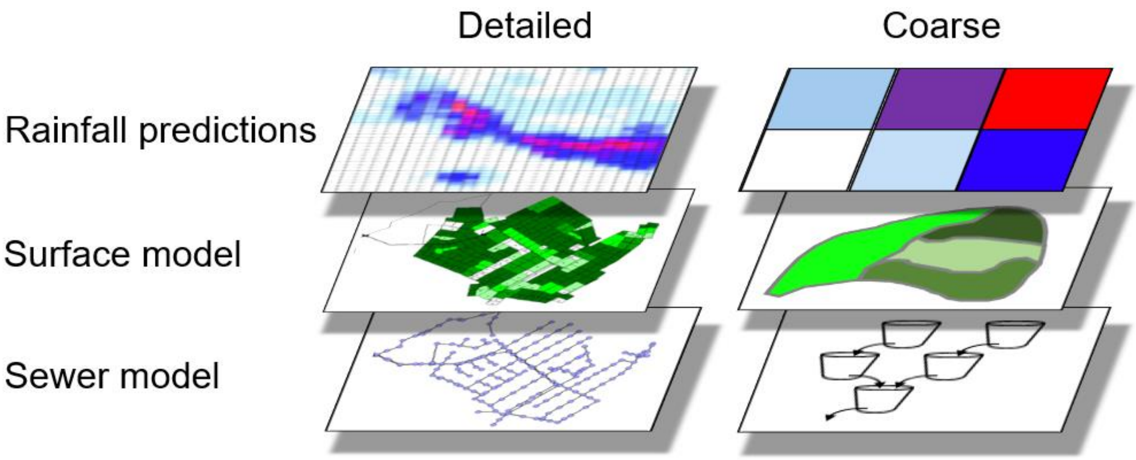


Figure 11: Consecutive layers/elements in a chain of forecasting models for urban runoff predictions (rows). Rainfall data, hydrological surface model and hydrodynamic sewer models of high detail intuitively fit together (left column), as do their coarse and simplified equivalents (right column).

Based on the literature and the results of this thesis, the following three approaches are ways to handle at least some of the mismatch issues of NWP-based forecasting and warning systems for small urban catchments:

1. For high-impact weather it is possible to only simulate one or a few worst-case scenarios and make decisions based on that. This could be the worst-case member in an ensemble or a high percentile such as the 95th in a probabilistic forecast (Jasper-Tönnies et al., 2018; René et al., 2018). It could also be a time series from the grid box within a neighborhood where the most severe rainfall intensities are predicted (Courdent et al., 2017, 2018; Olsson et al., 2017; Thorndahl et al., 2016). These attempts can keep a detailed hydrodynamic model as part of the forecasting setup, and take the approach of issuing a warning if the worst possible scenario exceeds a threshold, with the downside that they cannot provide probabilistic forecasts.
2. Others abandon the idea of a computationally expensive runoff model, and instead choose to use simple, conceptual models, which in turn can provide hundreds or thousands of simulations in real-time (Courdent et al., 2018; **Paper III**). A related approach would be to use NWP as input to fast surrogate models (also called emulators or meta-models) that have been trained to emulate the behavior of detailed runoff models at greatly reduced simulation times. Several surrogate models have been developed in recent years for e.g. flow (Lund et al., 2019; Thrysøe et al., 2019) and flood forecasting (Bermúdez et al., 2018). Machine learning models could be another option for computationally low-cost real-time predictions (Berkhahn et al., 2019).
3. Finally, it is possible to produce warnings and make decisions without simulating any runoff processes in real-time, and instead base actions on exceedance of predefined rainfall thresholds (Gaborit et al., 2013, 2016; Jasper-Tönnies et al., 2018; Yang et al., 2016, **Paper II**). With this approach runoff simulations can either not be used at all, or be made as offline, desktop exercises (e.g. with a detailed hydrodynamic model) that examine various potential flooding scenarios. These scenarios can then inform the choice of rainfall thresholds for a given warning or decision problem.

5 Conclusions

The advanced monitoring and modeling concepts, that are transforming urban drainage and wastewater systems from static solutions into actively managed assets, require high-quality input data. This thesis has made major strides towards operational use of two data sources that have highly desirable properties, but generally have been considered too unreliable and uncertain in the past: A-ISE sensors for monitoring NH_4^+, and NWP for rainfall forecasting.

The thesis showed that current standard operating procedures for how to maintain and operate A-ISE sensors at WRRFs need to be revisited. Recalibrating the sensors based on standard grab sample adjustments often led to an erratic and unreliable signal. A methodology was developed to improve the reliability of the A-ISE sensors by merging their signals with 24-h volume-proportional composite samples. These samples are widely available at WRRFs due to regulatory requirements, and implementation of the developed setup does thus not incur additional costs on operations. The merged dataset had much better data quality both for offline desktop studies and online estimates of NH_4^+. The reconstructed data was also more useful for training a software sensor. Since the use of composite samples is easy to implement and provides strong performance improvements, it should be pursued and further developed by water utilities.

A review of current NWP applications highlighted some key issues that must be taken into account by urban water practitioners. NWP products are designed with specific time scales in mind and should be used for urban drainage purposes that exist on similar scales. NWP skill is highly dependent on spatial scale and they are not expected to provide accurate prediction directly above a small urban catchment. This needs to be accounted for in how they are used. An often seen issue in the literature is use of hydrological models that are not suited for the NWP estimates that deliver their inputs.

Only a handful of urban water publications showcase real operational use of NWP while most studies so far have been desktop analyses. Many published studies do not possess much power in their analyses as they are based on evaluations of only a handful of rain events. This thesis based its results on large archives of historical forecasts, evaluated more than 100 rain events, and thus provide much more robust assessments. With this dataset it was possible to quantitatively evaluate the predictive performance of NWP rainfall estimates as a function of the weather type. Rain events with a lot of dynamic evolution and those consisting of small and scattered rain cells were the most difficult to predict for both NWP and radar nowcasts. A NWP product with improved initial conditions through assimilation of weather radar observations was able to compete with the more commonly used radar nowcasts. The NWP product had good skill in terms of operating a wet weather switch at a WRRF and was a vast improvement over a reactive control setup based solely on real-time observations of rainfall. While rainfall forecasts based on radar nowcasting drastically worsened with increasing forecast horizons, NWP was able to retain much of its skill for longer horizons.

NWP outputs are uncertain and post-processing methods can help describe and correct for this. The thesis investigated an intuitive post-processing method based on time-lagging of forecasts, which can account for some of the uncertainty related to specifying initial conditions. This method was able to highlight forecast consistency for the incoming weather situation and performed comparably to another post-processing method based on neighborhoods, which can account spatial uncertainty.

Overall, the thesis increased the reliability of A-ISE signals to the point where they may be used in real-time operations behind the primary clarifiers of WRRFs. It also assessed the potential of rainfall forecasts from NWP at small urban scales, and highlighted ways to properly employ them. More research and more real life experiences are still needed for improving the usefulness of A-ISE sensors and NWP, but the thesis showed that both technologies can have an important role to play in the increasingly digital operations of urban drainage and wastewater management.

6 Future research

6.1 Monitoring wastewater with A-ISE

The data merging technique for A-ISE sensors and composite samples should be further developed. Future tests will show which frequency of composite samples that are required for adequate data quality, and how that may vary between locations. The technique should be tested for other variables, such as nitrate, where outputs from low-cost sensors may be improved. The technique itself could be expanded to account for uncertainties during the merging process, e.g. through data assimilation techniques based on Kalman filters. This might also provide valuable uncertainty estimates on the measured values.

Further research should investigate which locations A-ISE sensors are suitable for with and without the merging technique. Within a WRRF, it would be interesting to test the technique for sensor locations before the primary clarifier and at the WWTP effluent where composites also often are available.

Obtaining raw potentiometric data from A-ISE sensors would allow for testing the assumption of relatively stationary slope parameters and the effect of interfering ions, which may be different for separate vs combined sewers.

6.2 Use of NWP for urban drainage purposes

Examination of the current literature has shown that there is a large need for further research and real case implementations before NWP-based warning and control schemes will be common within urban water management. The following points represent necessary and promising research avenues:

- **Large-sample studies**: Much of the literature consist of studies that examine a handful of rainfall events and thus do not possess statistical power. Several studies are inconclusive and finish with statements that call for further research on their hypotheses. There is a need for large-sample studies that examine NWP applications over longer periods, such as **Paper II** and **III** of this thesis. When analyses of forecasting setups are restricted to rain events only, they are missing the component of whether their systems produce significant amounts of false alarms in dry periods. Large samples are difficult to obtain for flood events in individual catchments, and more flood-oriented studies should rely on multi-catchment analyses for more robust findings.

- **Real-life case implementations**: Many studies, including those in this thesis, are post-hoc desktop investigations of how NWP-based decisions would have performed had the systems existed in the past. There is a need for more reports on actual forecasting systems and the problems and potential solutions they encounter to assist others in real-life implementations. This is especially true for RTC applications as the collective literature contains just one 7-day test at a real case (Stentoft et al., 2020) and five simulation studies: two simulation studies that mostly fail to get any benefit from NWP (Gaborit et al., 2013, 2016), a methodological study (Courdent et al., 2017), a small "proof-of-concept" paper (Courdent et al., 2015), and **Paper II**'s WRRF-oriented wet weather switching problem. Some promising avenues are control of large storage volumes such as urban tunnel systems that have been constructed in several cities (Palmitessa et al., 2018, 2021) and forecast-based optimization of energy use.
- **Proportionality and construction of the forecast chain**: As discussed above, several studies go awry in how they combine NWP products with hydrological and hydrodynamic models for various purposes. The discussion of this thesis suggests proportionality between the levels of detail in each component of the forecasting chain as an intuitive start, but there is a need for investigating which combinations of NWP products, post-processing methods, runoff models, and control algorithms that provide useful results for which purposes.
- **Post-processing**: Post-processing of raw deterministic and ensemble NWP outputs are essential for many purposes. National weather services have many users of their products and are therefore often not interested in tailoring post-processing schemes for individual end users. The urban drainage field will therefore need to take ownership of the developments of post-processing methods that are relevant for it. Simple, intuitive post-processing methods, such as neighborhoods and time-lagging, should continue to be refined. Statistical post-processing is becoming increasingly attractive due to the recent surge of open forecast data that allow for archiving and merging of multiple forecast products. This will also create opportunities for developing methods that turn pseudo-probabilistic ensemble products into actual probabilistic forecasts for improved decision-making. Much more research is needed for this as precipitation is one of the more difficult variables to predict and post-process (Vannitsem et al., 2020).

7 References

Albino, V., Berardi, U., Dangelico, R.M., 2015. Smart cities: Definitions, dimensions, performance, and initiatives. J. Urban Technol. 22, 3–21. https://doi.org/10.1080/10630732.2014.942092

Alferes, J., Tik, S., Copp, J., Vanrolleghem, P.A., 2013. Advanced monitoring of water systems using in situ measurement stations: data validation and fault detection. Water Sci. Technol. 68, 1022–1030. https://doi.org/10.2166/wst.2013.302

Alves, A., Gómez, J.P., Vojinovic, Z., Sánchez, A., Weesakul, S., 2018. Combining co-benefits and stakeholders perceptions into green infrastructure selection for flood risk reduction. Environ. - MDPI 5, 1–23. https://doi.org/10.3390/environments5020029

Åmand, L., Carlsson, B., 2012. Optimal aeration control in a nitrifying activated sludge process. Water Res. 46, 2101–2110. https://doi.org/10.1016/j.watres.2012.01.023

Atzori, L., Iera, A., Morabito, G., 2010. The Internet of Things: A survey. Comput. Networks 54, 2787–2805. https://doi.org/10.1016/j.comnet.2010.05.010

Autiosalo, J., Vepsalainen, J., Viitala, R., Tammi, K., 2019. A feature-based framework for structuring industrial digital twins. IEEE Access 8, 1193–1208. https://doi.org/10.1109/ACCESS.2019.2950507

Ballard, S.P., Li, Z., Simonin, D., Caron, J.F., 2016. Performance of 4D-Var NWP-based nowcasting of precipitation at the Met Office for summer 2012. Q. J. R. Meteorol. Soc. 142, 472–487. https://doi.org/10.1002/qj.2665

Bauer, P., Thorpe, A., Brunet, G., 2015. The quiet revolution of numerical weather prediction. Nature 525, 47–55. https://doi.org/10.1038/nature14956

Ben Bouallègue, Z., Theis, S.E., Gebhardt, C., 2013. Enhancing COSMO-DE ensemble forecasts by inexpensive techniques. Meteorol. Zeitschrift 22, 49–59. https://doi.org/10.1127/0941-2948/2013/0374

Berkhahn, S., Fuchs, L., Neuweiler, I., 2019. An ensemble neural network model for real-time prediction of urban floods. J. Hydrol. 575, 743–754. https://doi.org/10.1016/j.jhydrol.2019.05.066

Bermúdez, M., Ntegeka, V., Wolfs, V., Willems, P., 2018. Development and comparison of two fast surrogate models for urban pluvial flood simulations. Water Resour. Manag. 32, 2801–2815. https://doi.org/10.1007/s11269-018-1959-8

Blumensaat, F., Leitão, J.P., Ort, C., Rieckermann, J., Scheidegger, A., Vanrolleghem, P.A., Villez, K., 2019. How urban storm- and wastewater management prepares for emerging opportunities and threats: Digital transformation, ubiquitous sensing, new data sources, and beyond - A horizon scan. Environ. Sci. Technol. 53, 8488–8498. https://doi.org/10.1021/acs.est.8b06481

Bowler, N.E., Pierce, C.E., Seed, A.W., 2006. STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. Q. J. R. Meteorol. Soc. 132, 2127–2155. https://doi.org/10.1256/qj.04.100

Brendel, C.E., Dymond, R.L., Aguilar, M.F., 2020. Integration of quantitative precipitation forecasts with real-time hydrology and hydraulics modeling towards probabilistic forecasting of urban flooding. Environ. Model. Softw. 134, 104864. https://doi.org/10.1016/j.envsoft.2020.104864

Brok, N.B., Munk-Nielsen, T., Madsen, H., Stentoft, P.A., 2020. Unlocking energy flexibility of municipal wastewater aeration using predictive control to exploit price differences in power markets. Appl. Energy 280, 115965. https://doi.org/10.1016/j.apenergy.2020.115965

Bryan, G.H., Wyngaard, J.C., Fritsch, J.M., 2003. Resolution requirements for the simulation of deep moist convection. Mon. Weather Rev. 131, 2394–2416. https://doi.org/10.1175/1520-0493(2003)131<2394:RRFTSO>2.0.CO;2

Buizza, R., Houtekamer, P.L., Pellerin, G., Toth, Z., Zhu, Y., Wei, M., 2005. A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems. Mon. Weather Rev. 133, 1076–1097. https://doi.org/10.1175/MWR2905.1

Cecconi, F., Reifsnyder, S., Ito, Y., Jimenez, M., Sobhani, R., Rosso, D., 2019. ISE-ammonium sensors in WRRFs: Field assessment of their influencing factors. Environ. Sci. Water Res. Technol. 5, 737–746. https://doi.org/10.1039/c8ew00763b

Cecconi, F., Reifsnyder, S., Sobhani, R., Cisquella-Serra, A., Madou, M., Rosso, D., 2020. Functional behaviour and microscopic analysis of ammonium sensors subject to fouling in activated sludge processes. Environ. Sci. Water Res. Technol. 6, 2723–2733. https://doi.org/10.1039/d0ew00359j

Ciach, G.J., 2003. Local random errors in tipping-bucket rain gauge measurements. J. Atmos. Ocean. Technol. 20, 752–759. https://doi.org/10.1175/1520-0426(2003)20<752:LREITB>2.0.CO;2

Clark, P., Roberts, N., Lean, H., Ballard, S.P., Charlton-Perez, C., 2016. Convection-permitting models: A step-change in rainfall forecasting. Meteorol. Appl. 23, 165–181. https://doi.org/10.1002/met.1538

Courdent, V., Grum, M., Mikkelsen, P.S., 2018. Distinguishing high and low flow domains in urban drainage systems 2 days ahead using numerical weather prediction ensembles. J. Hydrol. 556, 1013–1025. https://doi.org/10.1016/j.jhydrol.2016.08.015

Courdent, V., Grum, M., Munk-Nielsen, T., Mikkelsen, P.S., 2017. A gain–loss framework based on ensemble flow forecasts to switch the urban drainage–wastewater system management towards energy optimization during dry periods. Hydrol. Earth Syst. Sci. 21, 2531–2544. https://doi.org/10.5194/hess-21-2531-2017

Courdent, V., Vezzaro, L., Mikkelsen, P.S., Mollerup, A.L., Grum, M., 2015. Using ensemble weather forecast in a risk based real time optimization of urban drainage systems. La Houille Blanche 101–107. https://doi.org/10.1051/lhb/20150025

Courdent, V.A.T., 2017. Numerical weather prediction and relative economic value framework to improve integrated urban drainage-wastewater management. Technical University of Denmark.

Cristiano, E., ten Veldhius, M.-C., van de Giesen, N., 2017. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review. Hydrol. Earth Syst. Sci. 21, 3859–3878. https://doi.org/10.5194/hess-21-3859-2017

Davis, J., Goadrich, M., 2006. The relationship between Precision-Recall and ROC curves. Proc. 23rd Int. Conf. Mach. Learn. - ICML '06 233–240. https://doi.org/10.1145/1143844.1143874

De Mulder, C., Flameling, T., Weijers, S., Amerlinck, Y., Nopens, I., 2018. An open software package for data reconciliation and gap filling in preparation of Water and Resource Recovery Facility modeling. Environ. Model. Softw. https://doi.org/10.1016/j.envsoft.2018.05.015

Du, J., Berner, J., Buizza, R., Charron, M., Houtekamer, P., Hou, D., Isidora, J., Mu, M., Wang, X., Wei, M., Yuan, H., 2018. Ensemble Methods for Meteorological Predictions, NCEP Office Notes 493. https://doi.org/https://doi.org/10.7289/V5/ON-NCEP-493

Ebert, E.E., 2009. Neighborhood Verification: A Strategy for Rewarding Close Forecasts. Weather Forecast. 24, 1498–1510. https://doi.org/10.1175/2009WAF2222251.1

Eggimann, S., Mutzner, L., Wani, O., Schneider, M.Y., Spuhler, D., Moy De Vitry, M., Beutler, P., Maurer, M., 2017. The potential of knowing more: A review of data-driven urban water management. Environ. Sci. Technol. 51, 2538–2553. https://doi.org/10.1021/acs.est.6b04267

Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.E., Quirmbach, M., Vaes, G., Vieux, B., 2004. Towards a roadmap for use of radar rainfall data in urban drainage. J. Hydrol. 299, 186–202. https://doi.org/10.1016/j.jhydrol.2004.08.004

Feddersen, H., 2009. Technical Report 09-14: A Short-Range Limited Area Ensemble Prediction System. Copenhagen, Denmark.

Ferriman, A., 2007. BMJ readers choose sanitation as greatest medical advance since 1840. BMJ 334, 111. https://doi.org/10.1136/bmj.39097.611806.DB

Gaborit, E., Anctil, F., Pelletier, G., Vanrolleghem, P.A., 2016. Exploring forecast-based management strategies for stormwater detention ponds. Urban Water J. 13, 841–851. https://doi.org/10.1080/1573062X.2015.1057172

Gaborit, E., Muschalla, D., Vallet, B., Vanrolleghem, P.A., Anctil, F., 2013. Improving the performance of stormwater detention basins by real-time control using rainfall forecasts. Urban Water J. 10, 230–246. https://doi.org/10.1080/1573062X.2012.726229

García, L., Barreiro-Gomez, J., Escobar, E., Téllez, D., Quijano, N., Ocampo-Martinez, C., 2015. Modeling and real-time control of urban drainage systems: A review. Adv. Water Resour. 85, 120–132. https://doi.org/10.1016/j.advwatres.2015.08.007

Garrido-Baserba, M., Corominas, L., Cortés, U., Rosso, D., Poch, M., 2020. The fourth-revolution in the water sector encounters the digital revolution. Environ. Sci. Technol. 54, 4698–4705. https://doi.org/10.1021/acs.est.9b04251

Golding, B.W., 2009. Long lead time flood warnings: reality or fantasy? Meteorol. Appl. 16, 3–12. https://doi.org/10.1002/met

Goudenhoofdt, E., Delobbe, L., 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. Hydrol. Earth Syst. Sci. Discuss. 13, 195–203. https://doi.org/10.5194/hessd-5-2975-2008

Grant, S.B., Saphores, J.D., Feldman, D.L., Hamilton, A.J., Fletcher, T.D., Cook, P.L.M., Stewardson, M., Sanders, B.F., Levin, L.A., Ambrose, R.F., Deletic, A., Brown, R., Jiang, S.C., Rosso, D., Cooper, W.J., Marusic, I., 2012. Taking the “waste” out of “wastewater” for human water security and ecosystem sustainability. Science (80-.). 337, 681–686. https://doi.org/10.1126/science.1216852

Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., Lindskog, M., Barkmeijer, J., Dahlbom, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., Alexander, C., Dowell, D., Liu, S., Ikuta, Y., Fujita, T., 2018. Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. Q. J. R. Meteorol. Soc. 144, 1218–1256. https://doi.org/10.1002/qj.3179

Haimi, H., Mulas, M., Corona, F., Marsili-Libelli, S., Lindell, P., Heinonen, M., Vahala, R., 2016. Adaptive data-derived anomaly detection in the activated sludge process of a large-scale wastewater treatment plant. Eng. Appl. Artif. Intell. 52, 65–80. https://doi.org/10.1016/j.engappai.2016.02.003

Haimi, H., Mulas, M., Corona, F., Vahala, R., 2013. Data-derived soft-sensors for biological wastewater treatment plants: An overview. Environ. Model. Softw. 47, 88–107. https://doi.org/10.1016/j.envsoft.2013.05.009

Hill, D., Kerkez, B., Rasekh, A., Ostfeld, A., Minsker, B., Banks, M.K., 2014. Sensing and cyberinfrastructure for smarter water management: The promise and challenge of ubiquity. J. Water Resour. Plan. Manag. 140, 01814002. https://doi.org/10.1061/(asce)wr.1943-5452.0000449

Hutton, C.J., Kapelan, Z., Vamvakeridou-Lyroudia, L., Savić, D., 2014. Real-time data assimilation in urban rainfall-runoff models. Procedia Eng. 70, 843–852. https://doi.org/10.1016/j.proeng.2014.02.092

Jasper-Tönnies, A., Hellmers, S., Einfalt, T., Strehz, A., Fröhle, P., 2018. Ensembles of radar nowcasts and COSMO-DE-EPS for urban flood management. Water Sci. Technol. wst2018079. https://doi.org/10.2166/wst.2018.079

Kaelin, D., Rieger, L., Eugster, J., Rottermann, K., Bänninger, C., Siegrist, H., 2008. Potential of in-situ sensors with ion-selective electrodes for aeration control at wastewater treatment plants. Water Sci. Technol. 58, 629–637. https://doi.org/10.2166/wst.2008.433

Karpf, C., Krebs, P., 2011. Quantification of groundwater infiltration and surface water inflows in urban sewer networks based on a multiple model approach. Water Res. 45, 3129–3136. https://doi.org/10.1016/j.watres.2011.03.022

Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., Bedig, A., Kertesz, R., Braun, T., Cadwalader, O., Poresky, A., Pak, C., 2016. Smarter stormwater systems. Environ. Sci. Technol. 50, 7267–7273. https://doi.org/10.1021/acs.est.5b05870

Korsholm, U.S., Petersen, C., Sass, B.H., Nielsen, N.W., Jensen, D.G., Olsen, B.T., Gill, R., Vedel, H., 2015. A new approach for assimilation of 2D radar precipitation in a high-resolution NWP model. Meteorol. Appl. 22, 48–59. https://doi.org/10.1002/met.1466

Laakso, T., Kokkonen, T., Mellin, I., Vahala, R., 2018. Sewer condition prediction and analysis of explanatory factors. Water 10, 1–17. https://doi.org/10.3390/w10091239

Langeveld, J., Van Daal, P., Schilperoort, R., Nopens, I., Flameling, T., Weijers, S., 2017. Empirical sewer water quality model for generating influent data for WWTP modelling. Water (Switzerland) 9, 1–18. https://doi.org/10.3390/w9070492

Leigh, C., Alsibai, O., Hyndman, R.J., Kandanaarachchi, S., King, O.C., McGree, J.M., Neelamraju, C., Strauss, J., Talagala, P.D., Turner, R.D.R., Mengersen, K., Peterson, E.E., 2019. A framework for automated anomaly detection in high frequency water-quality data from in situ sensors. Sci. Total Environ. 664, 885–898. https://doi.org/10.1016/j.scitotenv.2019.02.085

Liguori, S., Rico-Ramirez, M.A., Schellart, A.N.A., Saul, A.J., 2012. Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments. Atmos. Res. 103, 80–95. https://doi.org/10.1016/j.atmosres.2011.05.004

Löwe, R., Urich, C., Sto. Domingo, N., Mark, O., Deletic, A., Arnbjerg-Nielsen, K., 2017. Assessment of urban pluvial flood risk and efficiency of adaptation options through simulations – A new generation of urban planning tools. J. Hydrol. 550, 355–367. https://doi.org/10.1016/j.jhydrol.2017.05.009

Löwe, R., Vezzaro, L., Mikkelsen, P.S., Grum, M., Madsen, H., 2016. Probabilistic runoff volume forecasting in risk-based optimization for RTC of urban drainage systems. Environ. Model. Softw. 80, 143–158. https://doi.org/10.1016/j.envsoft.2016.02.027

Lund, N.S.V., Falk, A.K.V., Borup, M., Madsen, H., Mikkelsen, P.S., 2018. Model predictive control of urban drainage systems: A review and perspective towards smart real-time water management. Crit. Rev. Environ. Sci. Technol. 0, 4–6. https://doi.org/10.1080/10643389.2018.1455484

Lund, N.S. V., Madsen, H., Mazzoleni, M., Solomatine, D., Borup, M., 2019. Assimilating flow and level data into an urban drainage surrogate model for forecasting flows and overflows. J. Environ. Manage. 248, 109052. https://doi.org/10.1016/j.jenvman.2019.05.110

Martin, C., Vanrolleghem, P.A., 2014. Analysing, completing, and generating influent data for WWTP modelling: A critical review. Environ. Model. Softw. 60, 188–201. https://doi.org/10.1016/j.envsoft.2014.05.008

Messner, J.W., Mayr, G.J., Wilks, D.S., Zeileis, A., 2014. Extending Extended Logistic Regression: Extended vs. Separate vs. Ordered vs. Censored. Mon. Weather Rev. 140515130920004. https://doi.org/10.1175/MWR-D-13-00355.1

Mittermaier, M.P., 2007. Improving short-range high-resolution model precipitation forecast skill using time-lagged ensembles. Q. J. R. … 133, 1487–1500. https://doi.org/10.1002/qj.135

Nerini, D., Foresti, F., Leuenberger, D., Robert, S., Germann, U., 2019. A reduced-space ensemble Kalman filter approach for flow-dependent integration of radar extrapolation nowcasts and NWP precipitation ensembles 987–1006. https://doi.org/10.1175/MWR-D-18-0258.1

Neumann, M.B., Rieckermann, J., Hug, T., Gujer, W., 2015. Adaptation in hindsight: Dynamics and drivers shaping urban wastewater systems. J. Environ. Manage. 151, 404–415. https://doi.org/10.1016/j.jenvman.2014.12.047

Newhart, K.B., Marks, C.A., Rauch-Williams, T., Cath, T.Y., Hering, A.S., 2020. Hybrid statistical-machine learning ammonia forecasting in continuous activated sludge treatment for improved process control. J. Water Process Eng. 37, 101389. https://doi.org/10.1016/j.jwpe.2020.101389

Ochoa-Rodriguez, S., Wang, L.P., Gires, A., Pina, R.D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., Van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., Ten Veldhuis, M.C., 2015. Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. J. Hydrol. 531, 389–407. https://doi.org/10.1016/j.jhydrol.2015.05.035

Ochoa-Rodriguez, S., Wang, L.P., Willems, P., Onof, C., 2019. A Review of Radar-Rain Gauge Data Merging Methods and Their Potential for Urban Hydrological Applications. Water Resour. Res. 55, 6356–6391. https://doi.org/10.1029/2018WR023332

Ohmura, K., Thürlimann, C.M., Kipf, M., Carbajal, J.P., Villez, K., 2019. Characterizing long-term wear and tear of ion-selective pH sensors. Water Sci. Technol. 3, 541–550. https://doi.org/10.2166/wst.2019.301

Olsson, J., Pers, B.C., Bengtsson, L., Pechlivanidis, I., Berg, P., Körnich, H., 2017. Distance-dependent depth-duration analysis in high-resolution hydro-meteorological ensemble forecasting: A case study in Malmö City, Sweden. Environ. Model. Softw. 93, 381–397. https://doi.org/10.1016/j.envsoft.2017.03.025

Ort, C., Lawrence, M.G., Rieckermann, J., Joss, A., 2010. Sampling for Pharmaceuticals and Personal Care Products (PPCPs) and illicit drugs in wastewater systems: Are your conclusions valid? A critical review. Environ. Sci. Technol. 44, 6024–6035. https://doi.org/10.1021/es100779n

Palmitessa, R., Borup, M., Mikkelsen, P.S., 2018. Urban tunnel systems for conveyance and storage of storm- and wastewater: features, classification, and modelling, in: Proceedings of the 11th International Conference on Urban Drainage Modelling. Palermo, Italy, pp. 251–254.

Palmitessa, R., Mikkelsen, P.S., Law, A.W.K., Borup, M., 2021. Data assimilation in hydrodynamic models for system-wide soft sensing and sensor validation for urban drainage tunnels. J. Hydroinformatics In press. https://doi.org/10.2166/hydro.2020.074

Panepinto, D., Fiore, S., Zappone, M., Genon, G., Meucci, L., 2016. Evaluation of the energy efficiency of a large wastewater treatment plant in Italy. Appl. Energy 161, 404–411. https://doi.org/10.1016/j.apenergy.2015.10.027

Papias, S., Masson, M., Pelletant, S., Prost-Boucle, S., Boutin, C., 2018. In situ continuous monitoring of nitrogen with ion-selective electrodes in a constructed wetland receiving treated wastewater: An operating protocol to obtain reliable data. Water Sci. Technol. 77, 1706–1713. https://doi.org/10.2166/wst.2018.052

Pappenberger, F., Cloke, H.L., Persson, A., Demeritt, D., 2011. On forecast (in)consistency in a hydro-meteorological chain: Curse or blessing? Hydrol. Earth Syst. Sci. 15, 2391–2400. https://doi.org/10.5194/hess-15-2391-2011

Pedersen, J.W., Lund, N.S. V, Borup, M., Löwe, R., Poulsen, T.S., Mikkelsen, P.S., Grum, M., 2016. Evaluation of maximum a posteriori estimation as data assimilation method for forecasting infiltration-inflow affected urban runoff with radar rainfall input. Water 8. https://doi.org/10.3390/w8090381

Peleg, N., Marra, F., Fatichi, S., Paschalis, A., Molnar, P., Burlando, P., 2018. Spatial variability of extreme rainfall at radar subpixel scale. J. Hydrol. 556, 922–933. https://doi.org/10.1016/j.jhydrol.2016.05.033

Pulkkinen, S., Nerini, D., Pérez Hortal, A.A., Velasco-Forero, C., Seed, A., Germann, U., Foresti, L., 2019. Pysteps: An open-source Python library for probabilistic precipitation nowcasting (v1.0). Geosci. Model Dev. 12, 4185–4219.

https://doi.org/10.5194/gmd-12-4185-2019

Raftery, A.E., Gneiting, T., Balabdaoui, F., Polakowski, M., 2005. Using Bayesian model averaging to calibrate forecast ensembles. Mon. Weather Rev. 133, 1155–1174. https://doi.org/10.1175/MWR2906.1

Rauch, W., Urich, C., Bach, P.M., Rogers, B.C., de Haan, F.J., Brown, R.R., Mair, M., McCarthy, D.T., Kleidorfer, M., Sitzenfrei, R., Deletic, A., 2017. Modelling transitions in urban water systems. Water Res. 126, 501–514. https://doi.org/10.1016/j.watres.2017.09.039

René, J.-R., Madsen, H., Mark, O., 2013. A methodology for probabilistic real-time forecasting – an urban case study. J. Hydroinformatics 15, 751. https://doi.org/10.2166/hydro.2012.031

René, J.R., Djordjević, S., Butler, D., Mark, O., Henonin, J., Eisum, N., Madsen, H., 2018. A real-time pluvial flood forecasting system for Castries, St. Lucia. J. Flood Risk Manag. 11, S269–S283. https://doi.org/10.1111/jfr3.12205

Rieger, L., Jones, R.M., Dold, P.L., Bott, C.B., 2014. Ammonia-Based Feedforward and Feedback Aeration Control in Activated Sludge Processes. Water Environ. Res. 86, 63–73. https://doi.org/10.2175/106143013x13596524516987

Rieger, L., Takács, I., Villez, K., Siegrist, H., Lessard, P., Vanrolleghem, P.A., Comeau, Y., 2010. Data reconciliation for wastewater treatment plant simulation studies - planning for high-quality data and typical sources of errors. Water Environ. Res. https://doi.org/10.2175/106143009x12529484815511

Roberts, N.M., Lean, H.W., 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. Mon. Weather Rev. 136, 78–97. https://doi.org/10.1175/2007MWR2123.1

Schellart, A., Liguori, S., Krämer, S., Saul, A., Rico-Ramirez, M., 2014. Comparing quantitative precipitation forecast methods for prediction of sewer flows in a small urban area. IAHS-AISH Publ. 59, 1418–1436. https://doi.org/10.1080/02626667.2014.920505

Scheuerer, M., 2014. Probabilistic quantitative precipitation forecasting using Ensemble Model Output Statistics. Q. J. R. Meteorol. Soc. 140, 1086–1096. https://doi.org/10.1002/qj.2183

Schilling, W., 1991. Rainfall data for urban hydrology: what do we need? Atmos. Res. 27, 5–21. https://doi.org/10.1016/0169-8095(91)90003-F

Schneider, M.Y., Carbajal, J.P., Furrer, V., Sterkele, B., Maurer, M., Villez, K., 2019. Beyond signal quality: The value of unmaintained pH, dissolved oxygen, and oxidation-reduction potential sensors for remote performance monitoring of on-site sequencing batch reactors. Water Res. 161, 639–651. https://doi.org/10.1016/j.watres.2019.06.007

Schwartz, C.S., Kain, J.S., Weiss, S.J., Xue, M., Bright, D.R., Kong, F., Thomas, K.W., Levit, J.J., Coniglio, M.C., Wandishin, M.S., 2010. Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. Weather Forecast. 25, 263–280. https://doi.org/10.1175/2009WAF2222267.1

Sedlak, D., 2014. Water 4.0: the past, present, and future of the world's most vital resource. Yale University Press.

Sharma, A.K., Guildal, T., Thomsen, H.A.R., Mikkelsen, P.S., Jacobsen, B.N., 2013. Aeration tank settling and real time control as a tool to improve the hydraulic capacity and treatment efficiency during wet weather: Results from 7 years' full-scale operational data. Water Sci. Technol. 67, 2169–2176. https://doi.org/10.2166/wst.2013.129

Solon, K., Flores-Alsina, X., Kazadi Mbamba, C., Ikumi, D., Volcke, E.I.P., Vaneeckhaute, C., Ekama, G., Vanrolleghem, P.A., Batstone, D.J., Gernaey, K. V., Jeppsson, U., 2017. Plant-wide modelling of phosphorus transformations in wastewater treatment systems: Impacts of control and operational strategies. Water Res. 113, 97–110. https://doi.org/10.1016/j.watres.2017.02.007

Stentoft, P.A., Munk-Nielsen, T., Vezzaro, L., Madsen, H., Mikkelsen, P.S., Møller, J.K., 2019. Towards model predictive control: Online predictions of ammonium and nitrate removal by using a stochastic ASM. Water Sci. Technol. 79, 51–62. https://doi.org/10.2166/wst.2018.527

Stentoft, P.A., Vezzaro, L., Mikkelsen, P.S., Grum, M., Munk-Nielsen, T., Tychsen, P., Madsen, H., Halvgaard, R., 2020. Integrated model predictive control of water resource recovery facilities and sewer systems in a smart grid: Example of full-scale implementation in Kolding. Water Sci. Technol. 81, 1766–1777. https://doi.org/10.2166/wst.2020.266

Sun, J., Xue, M., Wilson, J.W., Zawadzki, I., Ballard, S.P., Onvlee-Hooimeyer, J., Joe, P., Barker, D.M., Li, P.W., Golding, B., Xu, M., Pinto, J., 2014. Use of nwp for nowcasting convective precipitation: Recent progress and challenges. Bull. Am. Meteorol. Soc. 95, 409–426. https://doi.org/10.1175/BAMS-D-11-00263.1

Theis, S.E., Hense, A., Damrath, U., 2005. Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. Meteorol. Appl. 12, 257. https://doi.org/10.1017/S1350482705001763

Therrien, J.-D., Nicolaï, N., Vanrolleghem, P.A., 2020. A critical review of the data pipeline: how wastewater system operation flows from data to intelligence. Water Sci. Technol. 82, 2613–2634. https://doi.org/10.2166/wst.2020.393

Thomann, M., Rieger, L., Frommhold, S., Siegrist, H., Gujer, W., 2002. An efficient monitoring concept with control charts for on-line sensors. Water Sci. Technol. 46, 107–116. https://doi.org/10.2166/wst.2002.0563

Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J.E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M.R., Molnar, P., 2017. Weather radar rainfall data in urban hydrology. Hydrol. Earth Syst. Sci. 21, 1359–1380. https://doi.org/10.5194/hess-2016-517

Thorndahl, S., Nielsen, J.E., Jensen, D.G., 2016. Urban pluvial flood prediction: A case study evaluating radar rainfall nowcasts and numerical weather prediction models as model inputs. Water Sci. Technol. 74, 2599–2610. https://doi.org/10.2166/wst.2016.474

Thorndahl, S., Poulsen, T.S., Bøvith, T., Borup, M., Ahm, M., Nielsen, J.E., Grum, M., Rasmussen, M.R., Gill, R., Mikkelsen, P.S., 2013. Comparison of short-term rainfall forecasts for modelbased flow prediction in urban drainage systems. Water Sci. Technol. 68, 472–478. https://doi.org/10.2166/wst.2013.274

Thrysøe, C., Arnbjerg-Nielsen, K., Borup, M., 2019. Identifying fit-for-purpose lumped surrogate models for large urban drainage systems using GLUE. J. Hydrol. 568, 517–533. https://doi.org/10.1016/j.jhydrol.2018.11.005

van Loosdrecht, M.C.M., Brdjanovic, D., 2014. Anticipating the next century of wastewater treatment. Science (80-.). 344, 1452–1453. https://doi.org/10.1126/science.1255183

Vannitsem, S., Bremnes, J.B., Demaeyer, J., Evans, G.R., Flowerdew, J., Hemri, S., Lerch, S., Roberts, N., Theis, S., Atencia, A., Bouallègue, Z. Ben, Bhend, J., Dabernig, M., de Cruz, L., Hieta, L., Mestre, O., Moret, L., Plenković, I.O., Schmeits, M., Taillardat, M., van den Bergh, J., van Schaeybroeck, B., Whan, K., Ylhaisi, J., 2020. Statistical postprocessing for weather forecasts – Review, challenges and avenues in a big data world. arXiv 1–29. https://doi.org/10.1175/bams-d-19-0308.1

Vezzaro, L., Grum, M., 2014. A generalised Dynamic Overflow Risk Assessment (DORA) for Real Time Control of urban drainage systems. J. Hydrol. 515, 292–303. https://doi.org/10.1016/j.jhydrol.2014.05.019

Vezzaro, L., Pedersen, J.W., Larsen, L.H., Thirsing, C., Duus, L.B., Mikkelsen, P.S., 2020. Evaluating the performance of a simple phenomenological model for online forecasting of ammonium concentrations at WWTP inlets. Water Sci. Technol. 81, 109–120. https://doi.org/10.2166/wst.2020.085

Villarini, G., Mandapaka, P. V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. J. Geophys. Res. Atmos. 113, 1–12. https://doi.org/10.1029/2007JD009214

Winkler, S., Rieger, L., Saracevic, E., Pressl, A., Gruber, G., 2004. Application of ion-sensitive sensors in water quality monitoring. Water Sci. Technol. 50, 105–114. https://doi.org/10.2166/wst.2004.0678

Wright, L., Davidson, S., 2020. How to tell the difference between a model and a digital twin. Adv. Model. Simul. Eng. Sci. 7. https://doi.org/10.1186/s40323-020-00147-4

Yang, C., Daigger, G.T., Belia, E., Kerkez, B., 2020. Extracting useful signals from flawed sensor data: Developing hybrid data-driven approaches with physical factors. Water Res. 185, 116282. https://doi.org/10.1016/j.watres.2020.116282

Yang, T.H., Hwang, G. Do, Tsai, C.C., Ho, J.Y., 2016. Using rainfall thresholds and ensemble precipitation forecasts to issue and improve urban inundation alerts. Hydrol. Earth Syst. Sci. 20, 4731–4745. https://doi.org/10.5194/hess-20-4731-2016

Yoon, S.-S., 2019. Adaptive blending method of radar-based and Numerical Weather Prediction QPFs for urban flood forecasting. Remote Sens. 11, 642. https://doi.org/10.3390/rs11060642

Zhang, D., Lindholm, G., Ratnaweera, H., 2018. Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. J. Hydrol. 556, 409–418. https://doi.org/10.1016/j.jhydrol.2017.11.018

8 Papers

I **Pedersen, J. W.**, Larsen, L. H., Thirsing, C. & Vezzaro, L. (2020). Reconstruction of corrupted datasets from ammonium-ISE sensors at WRRFs through merging with daily composite sample. Water Research, 185, 116227. DOI: 10.1016/j.watres.2020.116227.

II **Pedersen, J. W.**, Vezzaro, L., Vedel, H., Thirsing, C., Madsen, H. & Mikkelsen, P. S. (2021). Comparison of high-resolution numerical weather predictions and radar extrapolation forecasts from an urban drainage perspective. Submitted.

III **Pedersen, J. W.**, Courdent, V. A. T., Vezzaro, L., Feddersen, H., Vedel, H., Madsen, H. & Mikkelsen, P. S. (2021). Evaluation of time-lagged numerical weather prediction ensembles for urban runoff forecasting with ROC and PR analysis. Manuscript.

In this online version of the thesis, **Paper I-III** are not included but can be obtained from electronic article databases e.g. via www.orbit.dtu.dk or on request from:

DTU Environment
Technical University of Denmark
Bygningstorvet, Building 115
2800 Kgs. Lyngby
Denmark

info@env.dtu.dk.

The Department of Environmental Engineering (DTU Environment) conducts science-based engineering research within three sections: Circularity & Environmental Impact, Climate & Monitoring, Water Technology & Processing. The department dates back to 1865, when Ludvig August Colding gave the first lecture on sanitary engineering.

Department of Environmental Engineering

Bygningstorvet, Building 115
2800 Kgs. Lyngby
Tel. +45 4525 1600
Fax +45 4593 2850

www.env.dtu.dk