Sewer orientated framework for ensemble-based chance-constrained model predictive control.

Svensen, Jan Lorenz; Niemann, Hans Henrik; Falk, Anne Katrine Vinther; Poulsen, Niels Kjølstad

Published in:
Advanced Control for Applications: Engineering and Industrial Systems

Link to article, DOI:
[10.1002/adc2.68](https://doi.org/10.1002/adc2.68)

Publication date:
2021

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Svensen, J. L., Niemann, H. H., Falk, A. K. V., & Poulsen, N. K. (2021). Sewer orientated framework for ensemble-based chance-constrained model predictive control. *Advanced Control for Applications: Engineering and Industrial Systems*, 3(4), [e68]. <https://doi.org/10.1002/adc2.68>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

ARTICLE TYPE

Sewer Orientated Framework for Ensemble-based Chance-Constrained MPC

Jan Lorenz Svensen^{*1} | Hans Henrik Niemann² | Anne Katrine Vinther Falk³ | Niels Kjølstad Poulsen¹

¹Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

²Department of Electrical Engineering, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

³DHI Denmark, 2970 Hørsholm, Denmark

Correspondence

*Jan Lorenz Svensen, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Asmussens Allé 303B, 2800 Kgs. Lyngby, Denmark. Email: jlsv@dtu.dk

Present Address

Asmussens Allé 303B, 2800 Kgs. Lyngby, Denmark.

Funding Information

This research was supported by the supported by Innovation Fond Denmark through the Water Smart City project (project 5157-00009B).

Abstract

In this work, we present a framework for ensemble-based (E) chance-constrained (CC) model predictive control (MPC) in sewer systems. The framework considers the availability of ensemble forecasts and the difficulties with propagation of distributions; through distribution estimation. Utilizing a case study of the sewer network of the city of Aarhus in Denmark, the performance of the ECC-MPC framework is evaluated through simulations. The evaluations were based on linear models of the case study and compare the ECC-MPC performance with the performance of CC-MPC. Based on the simulations, it was found that the ECC-MPC performed comparable to the performance of the CC-MPC, not only in the context of overflow and outflow but also with respect to behavior in response to changes in different aspects of forecast uncertainties. Regarding the aspects, it was found that expectation offset biases in the forecast were affecting the performance of the CC- and ECC-MPC the most. While other aspects only had a reduced effect on the performances, within the ranges tested. With the comparable performances, it was found that ECC-MPC would work as an alternative approach to CC-MPC.

KEYWORDS:

stochastic MPC; Combined Sewer Overflow; chance-constrained; Ensemble; Sewer system

1 | INTRODUCTION

In the last few decades, research into Model Predictive Control (MPC) has been applied to many types of systems with promising results^{1,2}. One application is sewer systems^{3–7}, which will be the focus of this work. In sewer systems, MPC have been applied for its predictive abilities, to include the information from rain forecasts into the control, as well as the ability to handle constraints.

For MPC applied to sewer systems, the used models are usually simplified to consist of networks of interconnected virtual tanks; lumped models of volume capacity representing segments of the overall system's network of pipes and tanks. The performance of MPC is depending on the accuracy of the used model and the forecasted disturbances to the system; the model and forecast uncertainty. In the case of sewer systems, model uncertainty comes from the simplified model not perfectly describing the real system, while forecast uncertainty originates in the rain forecast of the inflows to the system.

Within the MPC research there are many approaches on how to handle the presence of uncertainty^{8–16}; from the robust MPC¹⁶ methods aiming for the worst-case scenario, to the stochastic MPC^{8,15} methods aiming for statistic coverage of the uncertainty. There is robust tube-based MPC (T-MPC) methods¹⁶, which utilizing "tubes" to systematical describe all realizations of the

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ad2.68

uncertainty, such that the control is robust against the worst-case scenario of the system. The tubes are based on the knowledge about the boundaries of the uncertainties, and results in a conservative constriction of the original constraints.

Similarly, there is the stochastic approach known as scenario-based MPC (S-MPC)¹⁵, which utilizes an ensemble of N forecast scenarios to statistically approximate the robust approach; through an N -doubling of the size of the deterministic MPC program.

For this work, we will consider the method known as chance-constrained MPC^{9,10,11,13,16} (CC-MPC) from the group of stochastic MPC methods as the approach to handling uncertainty. The CC-MPC methods utilizes knowledge about the stochastic distribution of the uncertainties, to formulate each of the inequality constraints as probabilistic constraints with a chosen probability γ of being true. The knowledge about the constraint distributions is used to obtain deterministic versions of the probability constraints, utilizing the quantile functions of the constraint distributions.

The CC-MPC method were chosen over the other approaches given that, the probability γ of each constraint provides the CC-MPC with design-freedom for the control solution regarding the sought level of conservatism; a feature the robust T-MPC does not have. Similarly, CC-MPC were chosen over the S-MPC, given its conservation of the original program size; which quickly can turn S-MPC into a gigantic program. The CC-MPC were also considered a good fit for the sewer system, given that rain forecasts can be considered stochastic disturbances, describing the likelihood of different inflow scenarios to the sewers.

The past research of CC-MPC include focusing on ensuring feasibility^{13,16}, simplifying the formulation of fx. joint probabilistic constraints^{9,10}, and applicability to the case of sewer systems¹¹.

The reliance on knowing distributions and quantile functions gives the CC-MPC method some drawbacks, given that constraint distributions are obtained from knowledge of the distributions of the model and forecast uncertainties across the prediction horizon of the CC-MPC. For most types of systems and uncertainty distributions, propagation of a distribution throughout the system would change the initial type of the distribution, e.g. adding together the squares of different types of distribution. The quantile function of the resulting constraint distribution might then become impractical to compute. Exceptions to such distribution changes include linear systems with additive uncertainty with a normal distribution.

In the context of sewer systems, the usage of CC-MPC is further complicated by the rain forecasts rarely being provided as a distribution. It is more commonly provided as a fixed-sized ensemble of statistical scenarios of the rain's temporal and spatial development, from where the distribution has to be determined.

The main contribution of this work are the suggestion of a framework for applying CC-MPC to the general case of sewer systems. The proposed framework aims to handle the issues with the propagation and lack of available initial distribution, through estimation of the constraint distributions based on propagated ensemble forecasts, as illustrated in Fig. 1. In the research, another solution to this problem has previously been proposed by approximating the probabilistic constraints¹⁷, instead of focusing on the constraint distribution and estimation.

A performance evaluation of the proposed framework of ensemble-based CC-MPC (ECC-MPC) is given, based on simulations of the Aarhus case study¹⁸ from Denmark. The performance evaluation is done by comparison to the performance of a CC-MPC with known quantile functions (propagated distribution). The Aarhus case study is discussed through a design model for MPC; autogenerated from a MIKE URBAN Hi-Fi model by DHI^{18,19}.

For the performance evaluation, we will consider a linear system model, where only forecast uncertainty is included. The uncertainties in the evaluation will be normal distributed with varying bounds and biases between simulations. Given that sewer systems experience combined sewer overflows (CSO) from their weir structures, the CC-MPC formulations utilized in this work are based on our previous proposed revised formulation¹¹. As sewer systems are intrinsically stable and feasible both in the real world and the used mathematics, our MPC design in this paper will not include an explicit control law¹⁶.

In the following section, we will first present a brief formulation of the MPC and CC-MPC used in the work, followed by the formulation and a discussion of the suggested ECC-MPC framework. In the third section, we discuss the design models of the Aarhus case study. In the fourth section, we discuss the simulations of ECC-MPC and CC-MPC, and the results of the performance comparison. We end the paper with a conclusion of the paper's results.

1.1 | Notation

The notations utilized in this paper are as follows; the superscripts *in*, *out*, *u* and *cso* indicate the inflow, outflow, control flow, and weir overflow respectively, while the subscript k indicates the sample number and ΔT notes the sampling time. A bullet $\bullet$ represents a subset or set of a function's variables, while Bold font indicates vectors and $\bar{f}$ notes the maximum of a given function $f(x)$. For a stochastic variable X , the notation $X \sim F$ indicates that X is following a distribution F . While $E\{X\}$ and σ_X^2 note the expectation and variance of X respectively, and $Pr\{X \leq y\}$ and $\Phi_X(y)$ are the probability function and cumulative

FIGURE 1 Conceptual framework layout of Ensemble-based CC-MPC.

distribution function (CDF) of X respectively for a given value y . The superscript E is used for an ensemble of scenarios of a given variable x and function f , as x^E and f^E respectively.

2 | MPC AND THE FRAMEWORK

The general formulation of MPC for systems with overflows, such as sewer systems can be written as

$$J = \min_{\mathbf{u}} f(\mathbf{x}, \mathbf{u}, \mathbf{z}^{ref}, \mathbf{w}, \mathbf{q}^{cso}) \quad (1)$$

$$\mathbf{x}_{k+1} = \mathbf{h}_{proc}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k, \mathbf{q}_k^{cso}) \quad \forall k \geq 0 \quad (2)$$

$$q_{k,i}^{cso} = \begin{cases} t_{w,i}(T_i(\bullet)), & T_i(\bullet) \geq 0 \\ 0 & \end{cases} \quad \forall i \in \{1 : N_{cso}\} \quad (3)$$

$$\mathbf{h}_k(\mathbf{x}_k, \mathbf{w}_k, \mathbf{u}_k, \mathbf{q}_k^{cso}) = \mathbf{0} \quad \forall k \geq 0 \quad (4)$$

$$\mathbf{g}_k(\mathbf{x}_k, \mathbf{w}_k, \mathbf{u}_k, \mathbf{q}_k^{cso}) \leq \bar{\mathbf{g}}_k \quad \forall k \geq 0 \quad (5)$$

Where $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{w}$ corresponds to the states of the system, the control of the system, and the inflow disturbances into the system, fx. rain, over the prediction horizon. The states evolves according to the process function $\mathbf{h}_{proc}$ in (2). The overflows are given by $\mathbf{q}^{cso}$ and defined by some function $t_{w,i}$ and the switching function T_i , where each i th overflow corresponds to the overflow structure of the system.

The above formulation can be simplified by approximating the overflows $\mathbf{q}^{cso}$ as optimization variables through a minimization cost³⁶²⁰. Thus increasing the number of variables in the formulation, and removing the logical part given in (3); resulting in a program that is simpler to compute, perhaps even convex. If the process equation of (2) is further substituted into the cost of (1) and constraints of (4) and (5), the simplified program can be written as

$$J = \min_{\mathbf{u}, \mathbf{q}^{cso}} f_q(\mathbf{x}_0, \mathbf{u}, \mathbf{z}^{ref}, \mathbf{w}, \mathbf{q}^{cso}) \quad (6)$$

$$\tilde{\mathbf{h}}_k(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso}) = \mathbf{0} \quad \forall k \geq 0 \quad (7)$$

$$\tilde{\mathbf{g}}_k(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso}) \leq \bar{\mathbf{g}}_k \quad \forall k \geq 0 \quad (8)$$

$$\mathbf{0} \leq \mathbf{q}^{cso} \quad (9)$$

where $\mathbf{x}_0$ is the initial state of the system and the cost function f_q contains the initial cost function f and the substituted process $\mathbf{h}_{proc}$, plus an additional cost for the approximation by minimization. While $\tilde{\mathbf{h}}_k$ and $\tilde{\mathbf{g}}_k$ are the equality and inequality constraints with the substituted process equation.

In the case of the CC-MPC with the assumption of available knowledge about uncertainty distribution, we introduce probabilistic constraints into the MPC design. The revised formulation of CC-MPC given by (10)-(16), addresses the oddity of a

probabilistic constraint on an intrinsically feasible constraint¹¹; such as a weir flow.

$$J = \min_{\mathbf{u}, \mathbf{q}^{cso}, \mathbf{s}, \mathbf{c}} E\{f_q(\mathbf{x}_0, \mathbf{u}, \mathbf{z}^{ref}, \mathbf{w}, \mathbf{q}^{cso})\} + f_s(\mathbf{s}, \mathbf{c}) \quad (10)$$

$$E\{\tilde{\mathbf{h}}_k(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})\} = \mathbf{0} \quad \forall k \geq 0 \quad (11)$$

$$\Phi_{\tilde{g}_{k,i}(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})}^{-1}(\gamma_i) \leq \bar{g}_{k,i} + s_{k,i} \quad \forall k \geq 0, \forall i \notin \mathcal{N}^{cso} \quad (12)$$

$$E\{\tilde{g}_{k,i}(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})\} \leq \bar{g}_{k,i} \quad \forall k \geq 0, \forall i \in \mathcal{N}^{cso} \quad (13)$$

$$\Phi_{\tilde{T}_i(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})}^{-1}(\gamma_i) \leq 0 + c_{k,i} \quad \forall k \geq 0, \forall i \in \mathcal{N}^{cso} \quad (14)$$

$$0 \leq s_{k,i} \leq \Phi_{\tilde{g}_{k,i}(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})}^{-1}(\gamma_i) - E\{\tilde{g}_{k,i}(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})\} \quad (15)$$

$$\mathbf{0} \leq \mathbf{q}^{cso}, \mathbf{c} \quad (16)$$

where $\mathcal{N}^{cso}$ is the set of constraints defining an overflow in the system, and the added slack variables $\mathbf{s}$ and $\mathbf{c}$ preserves the feasibility of the original MPC with regards to expectation, see (15), but ensures the probability constraint holds if possible through minimization of an added cost f_s . The set $\mathcal{N}^{cso}$ is utilized to determine which constraints can be formulated as a probability constraint directly (12) or are formulated as an expectation constraint (13). Wherefore the latter case has the addition of a probability constraint (14) based on the process substituted switching function $\tilde{T}_i$ corresponding to the overflow defined by the given constraint.

For some distributions such as the normal distribution, the quantile function of a constraint can be standardized as below

$$\Phi_{\tilde{g}(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})}^{-1}(\gamma) = E\{\tilde{g}(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})\} + \sigma\{\tilde{g}(\mathbf{x}_0, \mathbf{w}, \mathbf{u}, \mathbf{q}^{cso})\}\Phi^{-1}(\gamma) \quad (17)$$

Allowing for simplification of the optimization program into a deterministic formulation. For example, when considering additive uncertainty, the variance-quantile function term of the right-hand-side in (17) becomes constant with respect to the optimization variables. Later, in the performance evaluation, we will utilize this to construct a CC-MPC with known quantile function.

2.1 | Framework for probabilistic constraints

The formulation of CC-MPC given above relies on knowing the expectations and γ -quantiles of the constraints; which in general can be difficult to have knowledge about. In the case of uncertainties with a constant nature, such as a fixed distribution; the expectations and quantile functions of the constraints could be determined once and for all, therefore providing the needed knowledge for the CC-MPC. An example of this could be some types of model uncertainties affecting the constraints.

If one cannot be certain of the nature of the uncertainties are unchanged, then even knowing the exact distribution of the uncertainties at each time might not be of much use. Given that when propagating uncertainty distributions throughout the constraints, the distribution might very well change to a different type of distribution. For examples when the constraint distribution results in a mix of different types of distributions. The resulting constraint distribution might not have a defined quantile function or one that is easy to determine.

Therefore it might be more effective to determine the constraint distribution through an estimation, based on propagating scenarios of the uncertainty rather than distribution formulas. Where the scenarios describe both temporal and spatial aspects of the uncertainties. If the distribution of the uncertainty is known, then an ensemble of such scenarios can be generated for the propagation. This further means that the knowledge of the distribution of the uncertainty is no longer a necessity if an ensemble of uncertainty scenarios is available from another source.

In general, the propagation of the uncertainty ensemble $\mathbf{w}^E$ is given by (18), where an ensemble of n scenarios are given for m sources of uncertainties $\mathbf{w}$, and mapped to an ensemble of n scenarios of the i th constraint as a function of deterministic parameters and variables

$$\mathbf{w}^E \in \mathbb{R}^{m \times n} \rightarrow g_i(\mathbf{x}_0, \mathbf{w}^E, \mathbf{u}, \mathbf{q}^{cso}) = g_i^E(\mathbf{x}_0, \mathbf{u}, \mathbf{q}^{cso}) \in \mathbb{R}^{1 \times n} \quad (18)$$

Given that the constraint ensemble is a function of the optimization variables, the ensemble suffers the same computational issues that the standard CC-MPC has with the quantile functions for general constraint functions: the distribution depends on the optimization.

Under the right conditions for the relationship between uncertainty and variables, these issues disappear with regards to certain stochastic properties of the constraints. To start with, if the relationship is multiplicable then it follows from lemma 1,

that the expectation and variance of the constraint are deterministic. Similarly, if the relationship is additive, then from lemma 2, it follows that the quantile function of the constraint is deterministic in addition to the expectation and variance of the constraint.

Lemma 1 (Multiplicative properties). If the scalar constraint function is given by (19) and $\mathbf{w}$ is following some distribution $\mathcal{F}$

$$g(\mathbf{u}, \mathbf{w}) = \sum_{i=0}^N g_{1,i}(\mathbf{w})g_{2,i}(\mathbf{u}), \quad \mathbf{w} \sim \mathcal{F}, \quad g_{i,j}(\mathbf{x}) : \mathbb{R}^{n_x \times 1} \rightarrow \mathbb{R}, \forall i, j \quad (19)$$

Then it follows that the expectation is given by

$$E\{g(\mathbf{u}, \mathbf{w})\} = \sum_{i=0}^N E\{g_{1,i}(\mathbf{w})\}g_{2,i}(\mathbf{u}) \quad (20)$$

$$\sigma^2\{g(\mathbf{u}, \mathbf{w})\} = \sum_{i=0}^N \sum_{j=0}^N \sigma^2\{g_{1,i}(\mathbf{w}), g_{1,j}(\mathbf{w})\}g_{2,i}^2(\mathbf{u})g_{2,j}^2(\mathbf{u}) \quad (21)$$

Lemma 2 (Additive properties). If the scalar constraint function is given by (22) and $\mathbf{w}$ is following some distribution $\mathcal{F}$

$$g(\mathbf{u}, \mathbf{w}) = g_1(\mathbf{w}) + g_2(\mathbf{u}), \quad \mathbf{w} \sim \mathcal{F}, \quad g_{i,j}(\mathbf{x}) : \mathbb{R}^{n_x \times 1} \rightarrow \mathbb{R}, \forall i, j \quad (22)$$

Then it follows that the quantile function, expectation, and variance is given by

$$\Phi_{g(\mathbf{u}, \mathbf{w})}^{-1}\{\gamma\} = \Phi_{g_1(\mathbf{w})}^{-1}\{\gamma\} + g_2(\mathbf{u}) \quad (23)$$

$$E\{g(\mathbf{u}, \mathbf{w})\} = E\{g_1(\mathbf{w})\} + g_2(\mathbf{u}) \quad (24)$$

$$\sigma^2\{g(\mathbf{u}, \mathbf{w})\} = \sigma^2\{g_1(\mathbf{w})\} \quad (25)$$

The lemmas above are also applicable to the constraint ensemble, in the sense that the estimation of the constraint distribution with respect to certain stochastic properties only needs to consider the constraint part g_1 that includes the uncertainties. This can simplify the estimation and the following MPC program, by the formulation becoming deterministic.

For this reason, the definition of our framework will assume the uncertainty to be additive, such that the estimated quantile function is independent of the optimization variables. The procedure of the ECC-MPC framework can be summarized as the following steps:

- Obtain an Ensemble forecast, $\mathbf{w}^E$
- Propagate the Ensemble to the constraints, $\mathbf{g}^E$
- Estimate the constraint distributions, $\Phi_{g_i(\mathbf{u}, \mathbf{w})}^{-1}$
- Solve the CC-MPC
- Apply control solution
- Repeat

The procedure is also illustrated in Fig.1. The first step is to either produce ensemble forecast from a known distribution of the future disturbances or receive the ensemble forecast from another external source, e.g. weather institutes. The second step in the procedure provides us with the ensemble of constraint scenarios we need for the third step, by taking each forecast scenario in the ensemble and propagate it through the entire system description.

With the constraint ensemble ready, the third step is to estimate the distributions and quantiles of the constraints. There are many ways to do such estimation, one of them is a Pearson's χ^2 goodness of fit test; that allows one to estimate the constraint distribution in terms of well-known distributions such as normal, uniform, gamma, and others. This method has the benefit of estimating the constraint distribution as the best fitting distribution with a known quantile function, ready to be applied to the constraint formulation. In the simulations, presented later in this work, this approach to the estimation of the constraint distributions is utilized.

In the fourth step, we are ready to solve the optimization program of CC-MPC, and apply the found solution to the controlled system in step five. Then the procedure repeats as to prepare for the next optimization of the CC-MPC.

2.1.1 | Benefits and limitations

The ECC-MPC framework described above has both benefits and limitations to its usage. The framework has the benefits of not being reliant on either the presence of ensemble forecast data from external sources or forecast distributions, but being usable when either one is available. With regard to the stochastic nature of the uncertainty, the framework only assumes the expectation and quantile functions can be estimated. This means the framework is benefiting from not depending on assumptions such as the uncertainties being independently distributed; in order to simplify the propagation throughout the system and its cross-correlations.

Given the ECC-MPC framework is based on an estimation approach, the choice of approach is a limiting factor in the quality of the estimation, and therefore in the resulting controller. This also means the computation time of the ECC framework is reliant on the choice of estimation approach, while the traditional CC-MPC method does not rely on estimations, but on the distribution propagation, which also can increase computation time. The computational burden of ECC-MPC and CC-MPC are therefore hard to evaluate generally, as they depend on estimation approach and complexities between forecast distribution types/system dynamics.

Furthermore, if one relies on externally sourced ensemble forecasts then the size of the ensemble is also a limiting factor of the estimation. Another limiting aspect of the estimations is the necessity of the constraint ensembles being independent of each other, in order for the estimation to be statistically valid. This means the ensembles of the ensemble forecast also has to be independently generated, regardless of their origin.

As mentioned earlier the ECC-MPC framework assumes the uncertainties and optimization variables to have an additive relationship. This limits the types of systems it can be applied to, excluding relationships such as multiplicative, logarithmic, and others. For the multiplicative case, the framework could be used if the standardization shown in (17) is reasonable to assume, utilizing lemma 1 to achieve a deterministic constraint.

2.1.2 | Thoughts on other MPC-types

The ECC-MPC framework can easily be redesigned for other types of MPCs, by changing the distribution estimation step. If the desired controller is a tube-based controller from the class of robust MPCs, then the estimation step could either be changed to only estimating the indirectly assumed uniform distribution of the constraints, or directly utilizing the worst-case from the available constraint scenarios; obtained from the propagation of the ensembles. Similarly, other types of MPC could refit the estimation step to suit the type of data they rely on.

3 | AARHUS SEWER SYSTEM AND MPC DESIGN

As mentioned earlier, the system utilized in the performance evaluation later in this work is based on the Marselisborg sewer system, specifically the segment of the system covering the city of Aarhus. Aarhus is a city covering an area of 91km^2 with a population of 280534 in 2020²¹. Given that a sewer system is a network of pipes and tanks described by the Saint-Venant equations^{22,23}, a simpler description is usually preferred for the MPC design; such as virtual tank models. In such models, complete sections of the system are lumped together and described by their combined volume storage capacity and their interactions with other sections^{24,4}. A virtual tank model description of the Aarhus sewer system can be seen in Fig. 2. The shown description was autogenerated by DHI from a high fidelity MIKE URBAN model¹⁸ and translated to a design of deterministic MPC. One of the design criteria in the autogeneration among others; is that each section has a controlled outflow, and is extended from there until another section is reached.

In this work, we will utilize this model as both the simulation model in the later performance evaluation, but also as the design model of the deterministic MPC to formulate our CC- and ECC-MPC. The design of the MPCs is based on a modular approach, such that each module corresponds to one of the tanks in Fig. 2, and connected likewise so. The downflow connection of each tank can be seen in Table 1 in the link column.

The dynamics of the tanks utilized in the model are given by

$$V_{k+1} = AV_k + Bq_k = V_k + \Delta T q_k \quad (26)$$

where q_k is the sum of flows in and out of the tank, at time k . In the design of MPC, a prediction horizon H_p is chosen as a set number of samples to predict into the future of the system. The dynamics above can be reformulated to cover the entire

FIGURE 2 A schematic of the Aarhus sewer system model based on an autogenerated virtual tank description. The colored segments correspond to different quarters of the city, with their names written in the segments. The apparent layout differentiates from the real geography of the city.

prediction horizon as given below

$$\mathbf{V} = \mathbf{M}_0 \mathbf{V}_0 + \mathbf{M} \mathbf{q}, \quad \mathbf{V} = [V_1, V_2, \dots, V_{H_p}]^T, \quad \mathbf{q} = [q_0, q_1, q_2, \dots, q_{H_p-1}]^T \quad (27)$$

$$\mathbf{M} = \begin{bmatrix} B & 0 & \dots \\ AB & B & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots \\ A^{H_p-1}B & A^{H_p-2}B & \dots & B \end{bmatrix} = \Delta T \begin{bmatrix} 1 & 0 & \dots \\ 1 & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots \\ 1 & 1 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{H_p \times H_p}, \quad \mathbf{M}_0 = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^{H_p} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^{H_p \times 1} \quad (28)$$

The inflows and outflows of the tanks can be quantified as the controlled output q^{out} , CSO q^{cso} , internal inflow q^{in} and disturbance inflow q^{din} . This gives the following predictive dynamics

$$\mathbf{V} = \mathbf{M}_0 \mathbf{V}_0 + \mathbf{M} \mathbf{q}^{in} + \mathbf{M} \mathbf{q}^{din} - \mathbf{M} \mathbf{q}^{out} - \mathbf{M} \mathbf{q}^{cso} \quad (29)$$

In the MPC design, the CSO is treated as an optimization variable^{3,6,20}, by a minimization cost on accumulated CSO volume $\mathbf{V}_k^{cso}$, defined as

$$\mathbf{V}_k^{cso} = \sum_{i=0}^k \Delta T \mathbf{q}_i^{cso} \Rightarrow \mathbf{V}^{cso} = \mathbf{M} \mathbf{q}^{cso} \quad (30)$$

The cost function of the MPC design can be defined as the sum of cost functions $J_{D,i}$ from each module as seen in the left equation of (31), with minimization with respect to $\mathbf{q}^{out}$ and $\mathbf{q}^{cso}$. Similarly, the right equation shows the cost function of the

Tank	type	$\bar{V}[m^3]$	$\bar{q}[\frac{m^3}{s}]$	ρ_Q	ρ_C	link	Tank	type	$\bar{V}[m^3]$	$\bar{q}[\frac{m^3}{s}]$	ρ_Q	ρ_C	link
1	regular sc	300	1.50	-0.01	5000	25	20	regular	13001	0.22	-0.01	5000	9
2	regular sc	20	0.60	10^{-4}	5000	26	21	regular	10001	0.40	-0.01	5000	15
3	backflow sc	350	1.00	-0.01	200	4	22	regular	3501	0.18	-0.01	5000	14
4	backflow sc	844	1.80	-0.01	200	1	23	regular sc	90	1.20	10^{-4}	5000	21
5	dynamic	200	0.50	10^{-4}	5000	27	24	regular	3001	0.09	-0.01	15000	12
6	regular sc	200	0.80	10^{-4}	5000	28	25	regular	6699	1.50	-0.01	5000	
7	regular	100	0.10	10^{-4}	0.02	38	26	dynamic	3503	0.10	-0.01	10000	3
8	dynamic	10	0.04	10^{-4}	5000	19	27	regular	200	0.20	10^{-4}	0.02	18
9	regular	10	1.00	10^{-4}	5000	2	28	regular	60	0.70	10^{-4}	5000	16
10	regular	10	0.20	10^{-4}	5000	19	29	regular	100	0.20	10^{-4}	0.02	13
11	regular	150	0.35	10^{-4}	5000	17	30	regular	85	0.10	10^{-4}	200	31
12	regular	10	0.10	10^{-4}	5000	31	31	backflow	200	0.20	10^{-4}	20	10
13	regular sc	10	0.80	10^{-4}	5000	4	32	dynamic	100	0.04	10^{-4}	200	8
14	regular	10	0.20	10^{-4}	5000	4	33	dynamic	32	0.50	10^{-4}	200	5
15	regular	10	0.40	10^{-4}	5000	6	34	dynamic	25	0.50	10^{-4}	200	27
16	regular	10	0.70	10^{-4}	5000	1	35	dynamic	12	0.50	10^{-4}	200	27
17	regular sc	70	0.80	10^{-4}	5000	29	36	regular	80	0.30	10^{-4}	200	2
18	regular sc	50	0.15	10^{-4}	200	19	37	regular	90	0.40	10^{-4}	200	2
19	regular sc	85	0.98	10^{-4}	5000	11	38	regular	10	0.10	10^{-4}	5000	11

TABLE 1 The system data of the Aarhus network¹⁹

CC- and ECC-MPC and the module costs $J_{CC,i}$; with the slack variables $\mathbf{c}$ and $\mathbf{s}$ discussed earlier as optimization variables.

$$J = \min_{\mathbf{u}, \mathbf{q}^{cso}} \sum_{i=1}^{N_s} J_{D,i}, \quad J = \min_{\mathbf{u}, \mathbf{q}^{cso}, \mathbf{c}, \mathbf{s}} \sum_{i=1}^{N_s} J_{CC,i} \quad (31)$$

The different modules in the model have different physical meanings and are therefore formulated differently. Each module can be classified as one of five types of tanks; regular tanks, dynamic tanks, and backflow tanks, as well as a strict control version of the regular and backflow tanks. For the deterministic MPC, the modular cost function can be generalized as in (32) with linear costs on the outflow and accumulated CSO volume.

$$J_{D,i} = \rho_{Q,i}^T \mathbf{q}_i^{out} + \frac{\rho_{C,i}^T}{H_p} M \mathbf{q}_i^{cso} \quad (32)$$

Where ρ_Q and ρ_C are uniform weight vectors on the cost terms. A few modules diverge from the above cost and will be discussed later. In the case of the CC- and ECC-MPC, the cost function is an extended version of the deterministic case as seen in (33), and is utilized both for the ECC-MPC and the distribution-based CC-MPC.

$$J_{CC,i} = \rho_{Q,i}^T \mathbf{q}_i^{out} + \frac{\rho_{C,i}^T}{H_p} M \mathbf{q}_i^{cso} + \rho_S^T (\mathbf{s}_i + \mathbf{c}_i) \quad (33)$$

In the remainder of the paper, we will refer to (32) and (33) as the base costs $J_{D,i}^{base}$ and $J_{CC,i}^{base}$ respectively.

The added cost ρ_S on the slack variables is the same for all slack variables across all modules, with a value of a 100. The cost on the outflow and CSO for each module can be observed in Table 1. The table also shows the maximum volume and outflow of each tank as well as their module type. It can be seen that the tanks with larger volumes ($300 + m^3$) have a slight cost for maximizing the outflow, while small tanks aim at minimizing its usage. By comparison to Fig. 2, it can be seen that CSO connecting to external receivers are weighted higher (5000+), while internal CSOs have lower weights (200-), in order to avoid pollution outside the sewer system.

For the formulation of the modules in the paragraphs below, a few definitions are suitable. In (34), a variable $\mathbf{d}_i$ is given representing the disturbance and initial volume terms of the predictive dynamics for the i th tank, together with its ensemble

counterpart $\mathbf{d}_i^E$.

$$\mathbf{d}_i = [M_0 \ M] \begin{bmatrix} V_{0,i} \\ \mathbf{q}_i^{din} \end{bmatrix}, \quad \mathbf{d}_i^E = [M_0 \ M] \begin{bmatrix} V_{0,i}^E \\ \mathbf{q}_i^{din,E} \end{bmatrix} \quad (34)$$

The i th switching function discussed earlier in (3) is given by (35) over the prediction horizon H_p , as the volume above the i th tank's limit.

$$\mathbf{T}_i = \mathbf{M}_0 V_{0,i} + M \mathbf{q}_i^{in} + M \mathbf{q}_i^{din} - M \mathbf{q}_i^{out} - M_I \mathbf{q}_i^{cso} - \bar{\mathbf{V}}_i, \quad M_I = M - \Delta T I_{H_p} \quad (35)$$

where I_{H_p} is the identity matrix of size H_p , making the k th prediction of the CSO term independent of the k th CSO. For the CC-MPC based on distribution knowledge, we assume it follows a standardizable distribution as in (17) and that the different uncertainties are independent. Under these assumptions, the predicted deviation σ_i of the i th module is given by

$$\sigma_i = \sqrt{M_0^2 \sigma_{V_{0,i}}^2 + M^2 \sigma_{\mathbf{q}_i^{din}}^2} \quad (36)$$

where both the squares and root are to be understood as element operations.

3.1 | Modules - regular and backflow type

For the i th module belonging to either the regular or backflow type, the definition is identical with the difference lying in connections between tanks. For the deterministic MPC, the module is defined by the constraints given by (37)-(39), describing the volume, outflow, and CSO limitations.

$$\mathbf{0} \leq \mathbf{V}_i \leq \bar{\mathbf{V}}_i \quad (37)$$

$$\mathbf{0} \leq \mathbf{q}_i^{out} \leq \bar{\mathbf{q}}_i^{out} \quad (38)$$

$$\mathbf{0} \leq \mathbf{q}_i^{cso} \quad (39)$$

For the distribution-based CC-MPC, the constraints in (40)-(44) define the module as according to section 2, with the volume and switching function being substituted by (29) and (35) respectively.

$$\mathbf{0} + \sigma_i \Phi^{-1}(\gamma_i) - \mathbf{s}_i \leq E\{\mathbf{d}_i\} + M \mathbf{q}_i^{in} - M \mathbf{q}_i^{out} - M \mathbf{q}_i^{cso} \leq \bar{\mathbf{V}}_i \quad (40)$$

$$\mathbf{0} \leq \mathbf{q}_i^{out} \leq \bar{\mathbf{q}}_i^{out} \quad (41)$$

$$\mathbf{0} \leq \mathbf{q}_i^{cso}, \mathbf{c}_i \quad (42)$$

$$E\{\mathbf{d}_i\} + M \mathbf{q}_i^{in} - M \mathbf{q}_i^{out} - M_I \mathbf{q}_i^{cso} \leq \bar{\mathbf{V}}_i - \sigma_i \Phi^{-1}(\gamma_i) + \mathbf{c}_i \quad (43)$$

$$\mathbf{0} \leq \mathbf{s}_i \leq \sigma_i \Phi^{-1}(\gamma_i) \quad (44)$$

For the ECC-MPC, the definition is given by (45)-(49), where the quantile terms $\Phi_{\mathbf{d}^E}^{-1}$ and $\Phi_{-\mathbf{d}^E}^{-1}$ needs to be estimated.

$$\mathbf{0} + \Phi_{-\mathbf{d}^E}^{-1}(\gamma_i) - \mathbf{s}_i \leq M \mathbf{q}_i^{in} - M \mathbf{q}_i^{out} - M \mathbf{q}_i^{cso} \leq \bar{\mathbf{V}}_i - E\{\mathbf{d}_i^E\} \quad (45)$$

$$\mathbf{0} \leq \mathbf{q}_i^{out} \leq \bar{\mathbf{q}}_i^{out} \quad (46)$$

$$\mathbf{0} \leq \mathbf{q}_i^{cso}, \mathbf{c}_i \quad (47)$$

$$M \mathbf{q}_i^{in} - M \mathbf{q}_i^{out} - M_I \mathbf{q}_i^{cso} \leq \bar{\mathbf{V}}_i - \Phi_{\mathbf{d}^E}^{-1}(\gamma_i) + \mathbf{c}_i \quad (48)$$

$$\mathbf{0} \leq \mathbf{s}_i \leq \Phi_{-\mathbf{d}^E}^{-1}(\gamma_i) - E\{-\mathbf{d}_i^E\} \quad (49)$$

3.2 | Modules - strict control regular and backflow type

The strict control versions of the regular and backflow modules are using the same formulation as their counterpart, but includes extra constraints on the controlled outflow of the tank, as seen in (50). These added limitations relate the outflow to the volume of the tank by the volume-flow coefficient β^{25} . The lower constraint is scaled by the parameter $a_{sc,i}$, limiting how far the outflow can be from the physical limit.

$$a_{sc,i} \beta_i \mathbf{V}_i \leq \mathbf{q}_i^{out} \leq \beta_i \mathbf{V}_i \quad (50)$$

For regular modules with strict control, the value of $a_{sc,i}$ is 0.8, while it is 0.95 for the corresponding backflow modules.

For the definition of the modules with respect to the CC- and ECC-MPC, the added constraints introduce two new slack variables, with more cost terms added to the base module cost $J_{CC,i}^{base}$ from (33), as seen in (51).

$$J_{CC,i} = J_{CC,i}^{base} + \rho_S^T (s_{1,i} + s_{2,i}) \quad (51)$$

The module definition of the distribution-based CC-MPC is similar to above, the same formulation as the non-strict counterpart with the added constraints given by (52)-(54)

$$a_{sc,i} \beta_i E\{\mathbf{V}_i\} + a_{sc,i} \beta_i \sigma_i \Phi^{-1}(\gamma_i) - s_{1,i} \leq \mathbf{q}_i^{out} \leq \beta_i E\{\mathbf{V}_i\} - \beta_i \sigma_i \Phi^{-1}(\gamma_i) + s_{2,i} \quad (52)$$

$$\mathbf{0} \leq s_{2,i} \leq \beta_i \sigma_i \Phi^{-1}(\gamma_i) \quad (53)$$

$$\mathbf{0} \leq s_{1,i} \leq a_{sc,i} \beta_i \sigma_i \Phi^{-1}(\gamma_i) \quad (54)$$

From the constraints in (55)-(57), the additional constraints to defined the module for ECC-MPC are given, with the quantile function being dependent on scaled ensembles of $\mathbf{d}^E$.

$$a_{sc,i} \beta_i M(\mathbf{q}_i^{in} - \mathbf{q}_i^{out} - \mathbf{q}_i^{cso}) + \Phi_{a_{sc,i} \beta_i \mathbf{d}_i^E}^{-1}(\gamma_i) - s_{1,i} \leq \mathbf{q}_i^{out} \leq \beta_i M(\mathbf{q}_i^{in} - \mathbf{q}_i^{out} - \mathbf{q}_i^{cso}) - \Phi_{-\beta_i \mathbf{d}_i^E}^{-1}(\gamma_i) + s_{2,i} \quad (55)$$

$$\mathbf{0} \leq s_{2,i} \leq \Phi_{-\beta_i \mathbf{d}_i^E}^{-1}(\gamma_i) - E\{-\beta_i \mathbf{d}_i^E\} \quad (56)$$

$$\mathbf{0} \leq s_{1,i} \leq \Phi_{a_{sc,i} \beta_i \mathbf{d}_i^E}^{-1}(\gamma_i) - E\{a_{sc,i} \beta_i \mathbf{d}_i^E\} \quad (57)$$

3.3 | Modules - dynamic type

For the modules representing the dynamic tanks no constraints are added, but the control constraint in (38) is replaced with (58), where the upper limitation now depends on the inflows to the tank.

$$\mathbf{0} \leq \mathbf{q}_i^{out} \leq \bar{\mathbf{q}}_i^{out} + \mathbf{q}_i^{in} + \mathbf{q}_i^{din} \quad (58)$$

Given the replacement relies on the uncertain $\mathbf{q}^{din}$, the definitions for the CC- and ECC-MPC introduce a new slack variable with a new cost term associated with it

$$J_{CC,i} = J_{CC,i}^{base} + \rho_S^T s_{3,i} \quad (59)$$

The definition for distribution-based CC-MPC is given by (60) and (61), where the prior is the replaced constraint and the later is an added constraint on the new slack variable.

$$\mathbf{0} \leq \mathbf{q}_i^{out} \leq \bar{\mathbf{q}}_i^{out} + \mathbf{q}_i^{in} + E\{\mathbf{q}_i^{din}\} + s_{3,i} - \sqrt{\sigma_{q_i^{din}}^2} \Phi^{-1}(\gamma_i) \quad (60)$$

$$\mathbf{0} \leq s_{3,i} \leq \sqrt{\sigma_{q_i^{din}}^2} \Phi^{-1}(\gamma_i) \quad (61)$$

similarly, the constraints given in (62) and (63), gives the definition for the case of the ECC-MPC.

$$\mathbf{0} \leq \mathbf{q}_i^{out} \leq \bar{\mathbf{q}}_i^{out} + \mathbf{q}_i^{in} - \Phi_{-\mathbf{q}_i^{din,E}}^{-1}(\gamma_i) + s_{3,i} \quad (62)$$

$$\mathbf{0} \leq s_{3,i} \leq \Phi_{-\mathbf{q}_i^{din,E}}^{-1}(\gamma_i) - E\{-\mathbf{q}_i^{din,E}\} \quad (63)$$

3.4 | Modules - connections and operation

With the modules defined, the interconnection between them and specific operation goals can be discussed. The connections come in two types; backflows and downstream connections.

3.4.1 | Backflows

Backflows of a tank i are flows going upstream to the previous tank j inside the outflow pipe. In the two types of backflow modules, this is formulated as the CSO of the backflow of tank i affecting the outflow constraint of tank j . As seen in (64) for the general case of (38), but similarly added to other module type variants.

$$\mathbf{0} \leq \mathbf{q}_j^{out} + \mathbf{q}_i^{cso} \leq \bar{\mathbf{q}}_j^{out} \quad (64)$$

Tank j	2	3	4	10	11	15	16	17
Delay(min.)	20	10	4	2	5	20	3	14

TABLE 2 Delays in the Aarhus network¹⁹

Tank	7	21	24	27	29	31
α_{fac}	0.2857	0.5	0.2571	0.5714	0.2	0.5714

TABLE 3 The flow inhibitor factors α_{fac} in the Aarhus network¹⁹

3.4.2 | Downstream Connections

For the downstream connections, one has to consider possible flow delays and multiple sources of outflows and internal CSOs. The delays in the system can be observed in Table 2 with the tank of origin. For the receiving tank i , the tank inflow is formulated as (65), where the sets Q_c and Q_o define which CSO and outflows connect to the tank. While the matrices M_Δ and M_δ define the delays of the j th outflow.

$$\mathbf{q}_i^{in} = \sum_{j \in Q_c} \mathbf{q}_j^{cso} + \sum_{j \in Q_o} (M_{\Delta,j} \mathbf{q}_j^{out} + M_{\delta,j} \mathbf{q}_{\delta,j}^{out}) \quad (65)$$

$$M_{\Delta,j} = \begin{bmatrix} \mathbf{0}_{N_{\delta,j}} & \mathbf{0}_{N_{\delta,j}} & \cdots \\ 1 - \delta_{q,j} & \delta_{q,j} & 0 \\ 0 & 1 - \delta_{q,j} & \delta_{q,j} \\ \vdots & \ddots & \ddots \\ 0 & 0 & 1 - \delta_{q,j} & \delta_{q,j} \end{bmatrix} \in \mathbb{R}^{H_p \times H_p}, \quad M_{\delta,j} = \begin{bmatrix} \delta_{q,j} & (1 - \delta_{q,j}) & 0 & \cdots \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \delta_{q,j} & (1 - \delta_{q,j}) \\ \mathbf{0}_{H_p - N_{\delta,j}} & \mathbf{0}_{H_p - N_{\delta,j}} & \cdots & \mathbf{0}_{H_p - N_{\delta,j}} \end{bmatrix} \in \mathbb{R}^{H_p \times N_{\delta,j} + 1} \quad (66)$$

where $N_{\delta,j}$ is the number of whole sample delays, $\delta_{q,j}$ is the remaining fraction of a whole sample, and $\mathbf{q}_{\delta,j}^{out}$ is the delayed flows from the last sample.

3.4.3 | Flow Inhibitors

A few of the tanks in the system has features affecting the outflow, which are not described by the general modules discussed so far. These tanks follow the constraints in (67) instead of (38). The values of the inhibition factor α_{fac} can be seen for each affected tank in Table 3.

$$\mathbf{0} \leq \alpha_{fac} \mathbf{q}_i^{out} \leq \overline{\mathbf{q}}_i^{out} \quad (67)$$

3.4.4 | Reference

For the operation of the sewer system, there is a desire to keep the outflow of tank 1 at a specific flow rate q^{ref} of $0.9 \text{ m}^3/\text{s}$, if possible. Therefore an extra cost and constraints are added to the module for tank 1, defined as

$$J_{D,1} = J_{D,1}^{base} + \boldsymbol{\rho}_R^T (\mathbf{r}_1 + \mathbf{r}_2) \quad (68)$$

$$-\mathbf{r}_1 \leq \mathbf{q}_1^{out} - q^{ref} \leq \mathbf{r}_2 \quad (69)$$

$$\mathbf{0} \leq \mathbf{s}_4, \mathbf{s}_5 \quad (70)$$

where $\boldsymbol{\rho}_R$ is uniformly valued at 0.01. For the CC-MPCs, the addition is the same.

FIGURE 3 weather scenarios, temporal.

3.4.5 | Spare Volume

For the system, it is desirable that the tanks are not empty. Therefore each tank description is extended with an operational lower limit to operate at, as given in (71) for a limit at some fraction α_b of the maximum volume.

$$J_{D,i} = J_{D,i}^{base} + \rho_B^T(\mathbf{r}_{3,i}) \quad (71)$$

$$\alpha_b \bar{\mathbf{V}}_i - \mathbf{r}_{3,i} \leq \mathbf{V}_i \quad (72)$$

$$\mathbf{0} \leq \mathbf{r}_{3,i} \quad (73)$$

where the weight ρ_B is uniformly valued at 0.01, and α_b at 0.1. For the CC- and ECC-MPC, this restriction of the lower constraint is similar to their own probabilistic restriction, and therefore this extra restriction is only considered for the expectation:

$$\alpha_b \bar{\mathbf{V}}_i - \mathbf{r}_{3,i} \leq E\{\mathbf{V}_i\} \quad (74)$$

4 | RESULTS-PERFORMANCE

For the evaluation of the ECC-MPC framework, we will compare its performance against the performance of the discussed distribution-based CC-MPC. The deterministic MPC is used as a baseline for the comparison, to give a performance reference. The MPCs will all use a prediction horizon H_p of 20 samples and a sampling time of 300 seconds, giving 100 minutes of prediction. The sampling time is chosen large enough to cover the computational time of the MPCs, while also covering the dynamics of the real sewer system. For the ECC- and CC-MPC, the confidence level γ of the quantile functions is used as 90% for all quantiles. The ECC-MPC is given an ensemble forecast of 50 independent rain scenarios, to use for the estimation. The initial volume V_0 of each tank is assumed to be known perfectly and is $1m^3$ at the start of each simulation. For the simulations, 9 sets of historical weather scenarios from the city of Aarhus were utilized. In Fig. 3, the temporal evolution of the rain events is shown, clarifying the duration of the event and its intensities in m^3/s . Likewise in Fig. 4, the spatial distribution of the rain inflow is shown, quantifying the amount of volume m^3 each tank has received per event.

FIGURE 4 weather scenarios, spatial.

Simulation Nr.	1	2	3	4	5	6	7	8	9	10	11	12
Uncertainty Bounds κ	0.5	0.5	0.5	0.5	0.5	0.1	0.25	0.75	1	0.5	0.5	0.5
Scaled Bias a	0.8	0.9	1	1.1	1.2	1	1	1	1	1	1	1
Offset Bias b	0	0	0	0	0	0	0	0	0	0.005	0.01	0.1

TABLE 4 The sizes of the different types of forecast uncertainty introduced in each of the 12 simulation scenarios.

4.1 | Distribution estimation approach

In the simulations, the estimation approach utilized by the ECC-MPC is the Pearson's χ^2 goodness of fit test mentioned earlier. In this work, only two distributions have been tested for during the simulations, Normal and Uniform distributions, but could easily be extended to other well-known distributions. The uniform distribution was used as the fallback strategy, in case of estimation failure. While each probabilistic constraint is scalar, the estimation procedure used has assumed that the quantile functions within one specific module follows the same type of distribution (normal, uniform, etc.), and has estimated them simultaneous.

4.2 | Test Scenarios

In order to evaluate the performance of the controllers, three different types of uncertainties were introduced into the rain forecast w utilized by the ECC- and CC-MPC, while the deterministic MPC is given the correct rain forecast w_0 . The three types of forecast uncertainty are:

- forecast's uncertainty bound, $w \in E\{w\} \pm 3\sigma_w$
- scaled bias of the forecast expectation, $E\{w\} = aw_0$
- offset bias of the forecast expectation, $E\{f\} = w_0 + b$

where the standard deviation term σ_w is given by $\sigma_w = \frac{\kappa}{3}w_0$. Each type of uncertainty was given a default setting, such that only one type of uncertainty was being varied from simulation to simulation. In Table 4, each used combination of the three uncertainty types is shown. The default values are $\kappa = 0.5$, $a = 1.0$ and $b = 0.0$ respectively. For the simulations, the uncertainty added to the rain forecast was truncated normal distributed, in accordance with the three types of uncertainties discussed. In Fig. 5, the total CSO per event is shown for the default simulation scenario. It can be seen that both ECC- and CC-MPC are comparable in performance for the default simulation. While they both perform worse than MPC for the middle-sized events, and comparable to for small and large events, corresponding to a low loaded system and a saturated system respectively.

FIGURE 5 Total CSO Volume

FIGURE 6 Total outflow with MPC as the baseline; for variations in uncertainty bounds.

4.3 | Performance: Uncertainty Bounds

By varying the bounds of the uncertainty on the utilized forecasts, the effects on the controller can be evaluated. This allows us to evaluate the reliance on the certainty of the forecast. Considering the outflow of the system, we can see from Fig. 6, that the ECC-MPC operate comparably similar to CC-MPC, while providing larger outflows than MPC for the middle-sized events. Furthermore, the effects of varying the bounds appear to have only a slight effect on the performances of both ECC- and CC-MPC, with outflow generally increases as bound is increased. When focusing on the performance in the context of CSO, we have both the total CSO volume and the external CSO volume to consider. Shown for both ECC- and CC-MPC in Fig. 7 and Fig. 8 respectively, as percentage improvement in comparison to MPC performance. From the external CSO, we can observe clear improvements for both CC-MPCs, with a slight trend for high bounds to perform better as the event size increases. The general performance between ECC- and CC-MPC is comparable but not identical. From the total CSO, it can be seen that the price of improving the external CSO is generally a great increase in the internal CSO, rerouting the water internally leading to the observed increase in the outflow. It can further be seen that the external CSO improvement is very reliant on the spatial/temporal distribution of the rain inflow, while the total CSO only seems to depend on the event size.

4.4 | Performance: Scaled Bias

In practical applications, biases are natural to expect in the expected forecast of the distribution or ensemble. The ECC- and CC-MPC's performances under a constant scaling bias is evaluated in Fig. 9 and Fig. 10, for the total CSO and external CSO respectively. It can be seen that the general performance across events matches the performance discussed previously, with improvement in external CSO volume, at the cost of increased total CSO. The ECC- and CC-MPC are still comparable in their performance. With regard to the scaling bias, the CSO improvement has no general trend, with the exception of events with smaller percentage improvements, who shows a slight benefit of overestimating the rain forecast. From Fig. 11, the performance in the outflow volume is shown. Again, the performances are similar between the CC-MPCs and with larger outflow than the MPC.

FIGURE 7 Total CSO improvement in percentages with MPC as the baseline, positive percentages representing a reduction in CSO.

FIGURE 8 External CSO improvement in percentages with MPC as the baseline, positive percentages representing a reduction in CSO.

4.5 | Performance: Offset Bias

Now considering the bias on the expectation of the forecast is a constant offset bias instead of a scaling bias. The total outflow volume with offset biases can be observed in Fig. 12. While the performance of the ECC-MPC are comparable to that of CC-MPC; they both have a clear reliance on the bias with the outflow increasing as the bias increases. The same reliance can be seen about the improvement in CSO from Fig. 13 and Fig. 14 for total and external CSO respectively. For the external CSO, the improvements show a general reliance on the bias, in the sense that the improvement drops as the bias increases; to the point of the largest bias results in a deterioration of performance.

4.6 | Reflections on the Results

In all of the simulations, the performance of both CC-MPC and ECC-MPC were comparable, but not identical to each other. For both MPC formulations, the generation of both total and external CSO volume were similar across all of the tested rain scenarios and uncertainty variations.

The performances of both ECC- and CC-MPC showed that constant offset bias is the predominant influence on their performance. While the performances appear to be less variant of the bound on the uncertainty and scaled biases. A reason for this can be that the size of the scale biases is within the uncertainty bound, and that the simulations on uncertainty bound utilizes an expectation equal to the correct forecast, only resulting in more conservative solutions. While the larger offset bias might shift the correct forecast outside the bounds of the forecasts of both the ECC- and CC-MPC.

An observed drawback of ECC-MPC is the computation time. The ECC-MPC were computationally slower than the CC-MPC with average computation times of 5.96 seconds over 0.68 seconds. As one would expect with the added estimations for the case of a linear system with additive normally distributed uncertainty, where analytical propagation of distribution is possible.

The comparable performances of the ECC-MPC with respect to CC-MPC, shows that the ECC-MPC works as an alternative approach to the CC-MPC approach, performance-wise. In the tested case of linear system with normal distributed uncertainties, the CC-MPC is still the preferred approach, due to the aspect of computation.

However, given that the CC-MPC requires analytical propagation of the uncertainty distributions to obtain the constraint distributions, that either has to be computed offline for constant distributions or online for varying distributions, the computational difficulties and added work effort increases for the CC-MPC when dealing with nonlinear system and/or non-Gaussian distributed uncertainties.

FIGURE 9 Total CSO improvement in percentages with MPC as the baseline, positive percentages representing a reduction in CSO.

FIGURE 10 External CSO improvement in percentages with MPC as the baseline, positive percentages representing a reduction in CSO.

With the two approaches shows comparable performance in the results, and ECC-MPC only propagates ensembles, which are practically computational indifferent of system complexity; this suggest that the ECC-MPC should provide a practical alternative to CC-MPC, when the system is more complex; such as nonlinear and non-Gaussian distributed.

Lastly a small but curious result can be observed in the general performance of both ECC- and CC-MPC w.r.t. the external CSO; The CC- and ECC-MPC perform better than the baseline MPC with a perfect forecast. The reason for this is a consequence of using a finite prediction horizon. It is due to the forecast not including the whole rain event, and is an intrinsic trade-off of all types of MPC; future information can change the optimum trajectory. Stated more practically, the MPC finds the optimal solution within the finite horizon, while CC-, and ECC- MPC finds more conservative solutions. This can result in an initial state of the next time step's optimization that is at times more favorable in the long term.

5 | CONCLUSIONS

In this paper, we have suggested a framework for handling propagation of uncertainty distributions in chance-constrained(CC) model predictive control (MPC) based on ensemble forecasts and distribution estimation. The ensemble-based CC-MPC framework (ECC-MPC) have been evaluated by simulations on a model of the sewer network of Aarhus, Denmark, and compared in performance with the performance of CC-MPC.

The simulations utilized linear models with normal distributed forecast uncertainty, to produce a setup where both CC-MPC and ECC-MPC could easily be computed. For this particular system setup, it was found that CC-MPC and the ECC-MPC were comparable in their performances, and even shown the same responses to different changes in aspects of the uncertainties, such as expectation biases and bounds. The aspect of uncertainty which had most effect on ECC- and CC-MPC were found to be offset biases, while they were less sensitive to changes in scaling biases and uncertainty bounds.

From the performance analysis, we can conclude that the ECC-MPC framework performs comparatively to the CC-MPC, with some of the tested scenarios being slightly improved; others slightly worsen, but generally behaving similarly across the scenarios. As expected, it was also found that the additional estimations of ECC-MPC made it computationally slower than CC-MPC, in the simulated setup. An additional result with respect to CSO was that both ECC- and CC-MPC outperformed the deterministic MPC with perfect forecasts, due to their more conservative nature.

FIGURE 11 Total outflow with MPC as the baseline; for variations in scaling bias.

FIGURE 12 Total outflow with MPC as the baseline; for variations in offset bias.

FIGURE 13 Total CSO improvement in percentages with MPC as the baseline, positive percentages representing a reduction in CSO.

FIGURE 14 External CSO improvement in percentages with MPC as the baseline, positive percentages representing a reduction in CSO.

With ECC-MPC providing a way to circumvent the computational difficulties of the analytical propagation of stochastic distributions seen in CC-MPC; the ECC-MPC is suited for more general systems, such as the case of non-Gaussian and/or non-linear control problems, where analytical derivations become increasingly difficult. Based on this and the performance analysis, we conclude that the ECC-MPC framework is a practical alternative approach to CC-MPC.

References

1. Lee JH. Model Predictive Control: Review of the Three Decades of Development. *International Journal of Control, Automation, and Systems* 2011; 9(3): 415-424.
2. Qin SJ, Badgwell TA. A survey of industrial model predictive control technology. *Control Engineering Practice* 2003; 11: 733-764.
3. Gelormino MS, Ricker NL. Model Predictive Control of a Combined Sewer System. *INT. J. CONTROL* 1994; 59(3): 793-816.
4. Ocampo-Martinez C. *Model Predictive Control of Wastewater Systems*. Advances in Industrial ControlSpringer . 2010.
5. Puig V, Cembrano G, Romera J, et al. Predictive optimal control of sewer networks using CORAL tool: application to Riera Blanca catchment in Barcelona. *Water Science & Technology* 2009; 60(4): 869-874.
6. Halvgaard R, Falk AKV. Water system overflow modeling for Model Predictive Control. In: Proceedings of the 12th IWA Specialised Conference on Instrumentation, Control and Automation. ; 2017.
7. Marinaki M, Papageorgiou M. *Optimal Real-time Control of Sewer Networks*. Advances in Industrial ControlSpringer . 2005.
8. Mesbah A. Stochastic Model Predictive Control: An overview and Perspectives for Future Research. *IEEE Control Systems Magazine* 2016: 30-44. doi: 10.1109/MCS.2016.2602087
9. Celik O, Abdulsamad H, Peters J. Chance-Constrained Trajectory Optimization forNon-linear Systems with Unknown Stochastic Dynamics. In: IEEE. IEEE; 2019; Macau, China: 6828-6833.
10. Grosso JM, Ocampo-Martinez C, Puig V, Joseph-Duran B. Chance-constrained model predictive control for drinking waternetworks. *Journal of Process Control* 2014; 24: 504-516.
11. Svensen JL, Niemann HH, Falk AKV, Poulsen NK. Chance-Constrained Model Predictive Control - A reformulated approach suitable for sewer networks. *Advanced Control for Applications* 2020. , Submitted.
12. Raso L, Schwanenberg D, Giesen v. dN, Overloop vP. Short-term optimal operation of water systems using ensemble forecasts. *Advances in Water Resources* 2014; 71: 200-208.
13. Evans M, Cannon M, Kouvaritakis B. Linear Stochastic MPC under finitely supported multiplicative uncertainty. In: Proc. of the 2012 American Control Conference. ; 2012.
14. Garatti S, Campi MC, Garatti S, Prandini M. The scenario approach for systems and control design. *Annual Reviews in Control* 2009; 33: 149-157.
15. Campi MC, Garatti S, Prandini M. *Scenario Optimization for MPC*: 445–463; Cham: Springer International Publishing . 2019
16. Kouvaritakis B, Cannon M. *Model Predictive Control - Classical, Robust and Stochastic*. Advances Textbooks in Control and Signal ProcessingSpringer . 2016
17. Chai R, Savvaris A, Tsourdos A, Chai S, Xia Y, Wang S. Solving Trajectory Optimization Problems in thePresence of Probabilistic Constraints. *IEEE TRANSACTIONS ON CYBERNETICS* 202; 50(10): 4332-4345.
18. Drejer LH, Halvgaard R. Water Smart Cities Milestone 5: Offline Model Predictive Control (MPC) performance results for the Marselisborg, Aarhus case study. tech. rep., DHI; : 2020.
19. Drejer L, DHI . private communication; 2020.
20. Svensen JL, Niemann HH, Poulsen NK. Model Predictive Control of Overflow in Sewer Networks: A comparison of two methods. In: Proc. of the 4th Int. Conf. on Control and Fault-Toleran Systems. ; 2019: 412-417.

21. Danmarks-Statistik . <https://www.statistikbanken.dk/BY1>; 2020. Accessed: 30-09-2020.
22. Butler D, Davies JW. *Urban Drainage*. Spon Press. third ed. 2011.
23. Chow VT, Maidment DR, Mays LW. *Applied Hydrology*. McGraw-Hill . 1988.
24. Garcia L, Barreiro J, Escobar E, F. Tellez NQ, Ocampo-Martinez C. Modelling and Real-Time Control fo Urban Drainage Systems: A Review. *Advances in Water Resources* 2015; 85: 120-132.
25. Singh VP. *Hydrologic Systems. Volume 1: Rainfall-runoff modelling*. Prentice Hall . 1988.