UPTEC W09 008
Examensarbete 30 hp
Februari 2009

Wireless sensors for detecting toxic disturbances in a sewage system - a feasibility study

Trådlösa sensorer för detektion av giftstörningar i ett ledningsnät för avloppsvatten

Johannes Nygren

Abstract

Wireless sensors for detecting toxic disturbances in a sewage system - a feasibility study

Johannes Nygren

Wireless sensor networks (WSNs) are promising for monitoring variables that are hard to access, which could depend on lack of access to the commercial electricity net, or delocalized properties of the variable, requiring several point measurements. This is because WSN units are cheap and easy to install, since they do not require wiring.

This work consists of a literature study of wireless sensor networks and some simulations in SimuLink regarding a possible application of such networks. The proposed application being simulated is toxic monitoring in a sewer pipe which enters a wastewater treatment plant. If the toxic concentrations violate a certain threshold, the incoming wastewater will enter a storage tank, which is emptied into the activated sludge basin slowly, keeping toxic concentrations under the threshold.

The aim of this work is to provide some preliminary design and configuration recommendations for WSNs for this particular application. This work suggests a configuration of two WSN nodes, each with a chemical sensor, one at the inlet and one a bit upstream. More nodes are shown to increase expected system longevity by decreasing the energy consumption of individual nodes, since lower sample frequencies are shown to give the same performance of the storage strategy, compared to a case with only one node at the plant inlet. The excess energy consumption from unsynchronized WSNs is also investigated. If the time offset difference between the nodes is 1 minute, the individual energy consumption was still smaller than the individual energy consumption of a one-node configuration at the inlet, according to simulations.

Key words: Wireless sensor networks, urban water systems, pipe flow, toxic supervision, activated sludge process.

Department of Information Technology, Uppsala University, Box 337, SE-751 05 Uppsala
ISSN 1401-5765

Referat

Trådlösa sensorer för detektion av giftstörningar i ett ledningsnät för avloppsvatten

Johannes Nygren

Trådlösa sensornätverk ("Wireless sensor networks", WSNs) är lovande i syfte att mäta variabler som är svåråtkomliga, vilket kan bero på otillgänglighet till det kommersiella energinätet, eller att variabeln har olokaliserade egenskaper vilket kräver flera punktmätningar. Detta beror på att WSN-enheter är billiga och lätta att installera, eftersom de inte kräver någon sladdbaserad koppling.

Detta arbete utgör en litteraturstudie om trådlösa sensornätverk samt en modelleringsoch simuleringsstudie i SimuLink rörande en möjlig tillämpning av sådana nätverk. Den föreslagna tillämpningen som simuleras är giftövervakning i ett ledningsnät som leder till ett avloppsreningsverk. Om giftkoncentrationen överskrider ett särskilt gränsvärde, så kommer det inkommande avloppsvattnet att hamna i en utjämningsbassäng. Vattnet i lagringsbassängen pumpas långsamt till den biologiska reningen i avloppsreningsverket så att giftkonconcentrationen hela tiden hålls under gränsvärdet.

Syftet med detta arbete är att tillhandahålla några preliminära design- och konfigurationsrekommendationer för trådlösa sensornätverk i denna specifika tillämpning. Detta arbete föreslår en konfiguration bestående av två WSN-noder, båda med varsin giftsensor, där en är vid ingången till avloppsreningsverket och den andra är en bit uppströms. En konfiguration med flera noder visar sig öka den förväntade livslängden för systemet genom att minska energikonsumtionen hos individuella noder, eftersom lägre samplingsfrekvenser ger samma prestanda hos den temporära förvaringsstrategin, jämfört med ett fall där nätverket bara utgörs av en nod vid inloppet. Överskottet av energikonsumtion på grund av osynkroniserade nätverk undersöks också. Om tidsförskjutningen mellan noderna är en minut blir den individuella energikonsumtionen fortfarande mindre än den individuella energikonsumtionen av en ennodskonfiguration vid inloppet, enligt simuleringar.

Nyckelord: Trådlösa sensornätverk, urbana vattensystem, flöde i rör, giftövervakning, aktivslamprocessen.

Institutionen för informationsteknologi, Uppsala universitet, Box 337, SE-751 05 Uppsala
ISSN 1401-5765

Preface

This project has been financially supported by WISENET and is the master thesis project for the degree of M.Sc. in Environmental and Aquatic Engineering at Uppsala University.

In particular, I want to thank Bengt Carlsson for his supervision and suggestions, both on articles and modeling issues. This was necessary to complete this project in a reasonable timeframe.

I also want to send my gratitude to my subject reviewer Alexander Medvedev for some valuable opinions and Björn Halvarsson for helping me with SimuLink whenever I needed it.

Uppsala, February 2009

Johannes Nygren

UPTEC W 09 008, ISSN 1401-5765
Printed at the Department of Earth Sciences, Geotryckeriet, Uppsala University, Uppsala 2009.

Populärvetenskaplig sammanfattning

Johannes Nygren

Trådlösa sensornätverk är ett kommunikativt nätverk av sk. noder, som mäter med sensorer, kommunicerar med varandra och skickar data med hjälp av en sändare. Den vanligaste typen av sändare är radiosändare, men det finns även akustiska sändare som är lämpliga att använda för ett sensornätverk under vatten. Sensorerna som används kan teoretiskt vara vilka tänkbara sensorer som helst; temperaturgivare, tryckgivare, kemiska givare etc. Ett viktigt önskemål är att sensorerna, liksom sändarna samt övriga komponenter som noderna utgör, ska dra så lite energi som möjligt eftersom noderna går på batterier. Även nodernas mjukvara bör vara energisnål, dvs. noderna bör programmeras på rimliga sätt så de inte drar energi i onödan, till exempel genom att stänga av mottagaren när det är troligt att inga meddelanden kommer.

Eftersom noderna är förhållandevis billiga och lätta att installera på grund av att de inte kräver någon strömförsörjning från det kommersiella energinätet, finns det många möjliga användningsområden för de trådlösa sensornätverken. Några exempel på tidigare tillämpningar är olika miljöövervakningsuppställningar i områden där det inte finns elförsörjning, exempelvis för uppskattningar av antalet individer i vilddjurspopulationer, samt för prognoser om vulkanutbrott. Det faktum att noderna är billiga och lätta att installera gör även att man kan använda flera noder samtidigt och få bra helhetsinformation tack vare många punktmätningar. Detta har varit användbart exempelvis för styrning av lufttemperatur och luftfuktighet inomhus, genom att ventilera enbart när det behövs.

Ett tänkbart tillämpningsområde är giftövervakning i ett ledningsnät för att kunna förvarna ett avloppsreningsverk om inkommande giftstötar. Vissa substanser har visat sig hämma den biologiska reningen som finns i de flesta avloppsreningsverk genom att slå ut floran av mikrorogansimer. Till exempel har nitrifierande mikroroganismer visat sig särsklit känsliga för gifter. Syftet med denna rapport är att undersöka genomförbarheten av en sådan tillämpning, beroende på hur långsamma sensorer man har tillgång till, var noderna är placerade samt vilken samplingstid som används.

Resultaten är uteslutande baserade på datorsimuleringar i en modell som är byggd i SimuLink. Vattenflödet i kloakerna modelleras med Mannings ekvation, och transporten av gift modelleras med en numerisk dispersionsalgoritm. Typen av gift är godtyckligt. I modellen antas att det bara finns ett gift, vars koncentration ges som $\mu g\ L^{-1}$. Om sensorerna registrerar giftkoncentrationer som överstiger ett visst tröskelvärde på 30 $\mu g\ L^{-1}$ skickar de data till avloppsreningsverket, som gör att inkommande avloppsvatten pumpas till en utjämningstank. När giftpulsen har

avlägsnat sig töms innehållet i förvaringstanken sakta in till aktivslambassängen under kontrollerade former, så att inte koncentrationen överstiger gränsvärdet. Tanken är att prestandan hos förvaringsstrategin endast ska bero på hur bra mätningar på giftkoncentration man har tillgång till.

Energikonsumtion hos noderna modelleras också. Syftet med det är att få en uppskattning om systemets hållbarhet, som bland annat är en funktion av energikonsumtionen hos individuella noder. Det visar sig att ett trådlöst nätverk på två noder, där en är vid inloppet till reningsverket och där den andra är 5 km uppströms ledningen, kan avlasta energikonsumtionen hos individuella noder avsevärt, jämfört med att bara ha en nod vid inloppet. Detta beror på att man kan välja en betydligt lägre samplingstid och ändå få samma prestanda hos förvaringsstrategin (dvs. strategi för att avleda avloppsvattnet till lagringsbassängen). Noden uppströms kan varna noden vid inloppet att den registrerar höga gifthalter med en larmsignal. Noden vid inloppet kan sedan öka samplingsfrekvensen (vilket är ekvivalent med att minska samplingstiden) ett tag efter larmsignalen.

Om nätverket är osynkroniserat visar sig energikonsumtionen hos noden uppströms öka betydligt, eftersom den måste skicka samma larmsignal flera gånger. Detta beror på att noden vid inloppet hade sin mottagare avstängd vid fel tillfälle, på grund av den dåliga tidssynkroniseringen. Tidssynkronisering verkar ha mer betydelse för energikonsumptionen än för prestandan hos förvaringsstrategin.

I det enklaste scenariot mäts inte flöden i ledningsnätet. Därför måste man använda tumregler för tidsförskjutningen från giftdetektion till dess att giftstöten når inloppet. Blir flödet oväntat stort, kommer alltså prestandan hos förvaringsstrategin att bli väldigt dålig, eftersom systemet inte hinner reagera innan giftkoncentrationen i aktivslambassängen skjutit i höjden. När flödesmätare införs blir prestandan bättre, men inte tillräckligt bra eftersom giftsensorerna har en inneboende långsamhet, vilket gör att de inte hinner med att detektera giftet innan koncentrationen i aktivslambassängen överstigit tröskelvärdet. Notera att giftkoncentrationen i aktivslambassängen stiger fortare om inflödet är stort. För att åtgärda detta, testas ett scenario där noden uppströms sänder data direkt till avloppsreningsverket istället för att larma noden nedströms. I det läget överstiger inte giftkoncentrationen i aktivslambassängen tröskelvärdet, vilket sker på bekostnad av väldigt hög energiåtgång för noden uppströms. Det kan vara nödvändigt att ha detta som en tillfällig åtgärd ifall det kommer en giftstöt under höga flöden om sensorerna är relativt långsamma.

Contents

1 Introduction **1**

2 Wireless sensor networks **2**
2.1 Communication 3
2.2 Localization 4
2.3 Time synchronization 5
2.4 Compression and aggregation 7
2.5 Automatic control using WSNs 8
2.6 Sensors in WSNs 10

3 Applications in Urban Water Systems **11**
3.1 Leak location in water pipe networks 11
3.2 Reduction of Combined Sewer Outflow (CSO) events 11

4 Toxic effects on the activated sludge process **13**

5 Methods **14**
5.1 The sewer pipe net model 15
5.1.1 Model ranges 22
5.2 The wastewater treatment plant model 24
5.3 The WSN Model 27
5.4 Scenarios and simulation 30
5.4.1 Performance tests with different sampling times 31
5.4.2 Realization of time synchronization 34
5.4.3 Flow measurements with upstream node compensation . . . 34
5.4.4 Example scenario 1: Constant flow with one sensor 35
5.4.5 Example scenario 2: Constant flow with two sensors 40
5.4.6 Implementation of main scenarios 45

6 Results and discussion **47**
6.1 Scenario 1: One node at the plant inlet 47
6.2 Scenario 2: Two nodes in pipe 3 50
6.3 Scenario 3: Two nodes in pipe 3 with variable flow 54

7 Conclusions **59**

References **60**
Personal communication 61

Glossary

Active sampling period	Asp	28
Area of threshold violation	A_v	32
Estimated detection time delay	$t_{est,D}$	44
Estimated inlet time delay	$t_{est,in}$	44
Flow from storage tank to activated sludge basin	Q_{ta}	24
Flow to activated sludge basin from incoming wastewater	Q_{ain}	24
Hydraulic retention time in activated sludge basin	t_{ret}	24
Incoming toxic concentration to pipe 1	c_1	22
Incoming toxic concentration to pipe 2	c_2	22
Incoming toxic concentration to wastewater treatment plant	c_{in}	24
Inflow to pipe 1	Q_1	22
Inflow to pipe 2	Q_2	22
Inflow to storage tank	Q_{tin}	25
Inflow to wastewater treatment plant	Q_{in}	24
Passive sampling period	Psp	28
Pipe segment length	l_p	19
Response time delay	t_r	42
Sampling period limit	$t_{s,lim}$	33
Sum of energy consumption during simulation	E_{tot}	33
Switch response time delay	$t_{r,sw}$	57
Threshold of toxic concentration	$c_{threshold}$	24
Threshold violation	T_{viol}	32
Time step length in pipe model	t_s	19
Toxic concentration in activated sludge basin	c_a	24
Toxic concentration in storage tank	c_t	24

1 Introduction

The interest for wireless sensor networks (WSNs) has increased rapidly in recent years due to their potential applications. Some scientists believe that wireless sensor networks will have a comparable impact on society as internet had. However, there seem to be few suggested applications for urban water systems in the literature. This report mentions two previous applications of WSNs in urban water systems. The first one is about locating leaks in water pipe networks with acoustic sensors, see [2] and [10]. The other is about reducing the risk of *combined sewer outflow* (CSO) events, which means overflow from sewer nets with combined wastewater and backwater, by utilizing the storage capacity in the combined sewer systems in a better way [9].

The purpose of this work is to propose a new way to apply WSNs to an urban water system, and simulate the potential benefits/drawbacks of this application, in terms of performance and energy consumption. Another purpose is to give some preliminary recommendations on relevant WSN aspects, such as time synchronization and sample frequencies.

The proposition of WSN application in this work is to monitor concentration levels of an arbitrary toxin which is inhibitory to the activated sludge process in a wastewater treatment plant. This is done using a small WSN equipped with chemical sensors, located in a sewer pipe net. The incoming wastewater switches over to a storage tank if the toxic concentrations become too high. The storage tank is then slowly emptied into the activated sludge basin, getting steady toxin levels.

The layout of this report is as follows. In the second chapter, some background information about WSNs is given. The chapter contains technical issues such as communication, network localization, time synchronization and control issues specific for WSNs. This is followed by some examples of applications in urban water systems in Chapter 3. Chapter 4 gives some background information on toxins and their inhibitory effects on the activated sludge process. Chapter 5 describes the model and two simulation scenarios, to describe how the model behaves in practice. It also describes the three main scenarios being simulated, whose results are presented in Chapter 6. The conclusions and recommendations for future work is discussed in Chapter 7.

2 Wireless sensor networks

A wireless sensor network consists of a network of so-called nodes. The nodes are in general battery driven individual units communicating with each other with radio transmitters. They are, in turn, connected to suitable sensors (which can be thermometers, pressure sensors, optical pulse sensors, etc). The nodes communicate to a sink node, either by direct communication or hop-by-hop communication depending on the location of the node. The sink node could, for example, be wired to a laptop, where the data is stored.

Wireless sensor networks have many beneficial applications, particularly since they are cheap and easy to install. Wireless monitoring systems are necessary if the monitored environment does not have installed infrastructure for energy. A typical example is monitoring of animal behavior or other animal population variables in their natural habitat.

Since the networks are wireless, the finite energy budget is a primary design constraint. Therefore, distributed signal processing is commonly used within the network matrix, to reduce the data amount transmitted between the nodes, since communications are a key energy consumer [3]. Processing of data in a distributed way can reduce communication cost, compared to the case where all raw data is sent to the sink node which performs all processing.

Besides energy economization, the WSNs face other technical challenges. The WSN must identify and adapt to resulting distributions of nodes, to simplify ad-hoc deployment of nodes, or if some nodes or sensors change position over time. Also, unattended operation requires self-configuration of the nodes. To address these problems, the following strategies are widely used [3]:

- Collaborative signal processing among nodes that have experienced a common stimulus.
- Exploiting redundancy of nodes in the system, by not letting more nodes work than what is necessary for coverage demand.
- Signal processing manufactured to minimize energy consumption.
- A hierarchical, tiered architecture where higher capacity elements can offload other elements when necessary.

Wireless sensor networks face many new technical challenges. In the following, some of these issues are discussed. One technical aspect which is not mentioned below is security. Security (from malicious agents) also faces new challenges for WSNs, since the nodes communicate through the whole medium instead of isolated wires, and since cryptation is a doubtful solution due to the excess communication

load. This problem, however, is not discussed since it is not the main topic of this work, see [17] for a further discussion.

2.1 Communication

Each sensor node in a WSN uses the protocol stack to communicate with one another and the sink node. Communication in WSNs can be divided into several layers with different functions. A layer is a set of protocols with specific communication tasks.

One of those layers is the *data-link layer*, which deals with communication between two nodes that share the same link. Medium Access Control (MAC) protocols, which is an important part of the data-link layer, tries to ensure that two nodes does not interfere with each other's transmissions, and deals with the situation when they do. There exist different MAC protocols with different ways of dealing with the problem. While traditional MAC protocols focus on maximizing package throughput and minimize latency, MAC protocols for WSNs focus on minimizing energy consumption.

The nodes commonly waste much energy by keeping their radios in receiver mode and listen for transmissions, since they do not know when a message is going to be sent to them. The T-MAC protocol, described in [15], uses an active/sleep duty cycle to reduce energy wasted on idle listening. During active mode, the node's radio can either be in receiver mode, or transmit messages themselves. During sleep mode, the node turns off it's radio to save energy. Messages are queued during sleep mode, and then bursted in active mode, rather than spreading them over a large active time interval. The nodes go back to sleep mode if a small, preset time period passes, without any message transfers taking place.

T-MAC performs just as good as a protocol with fixed duty cycles in simulations with homogeneous load, which is up to 98% reductions in energy consumption compared to the classic CSMA protocol (a protocol with no duty cycles). During variable load, T-MAC is five times better at conserving energy than the fixed duty cycle protocol, according to simulations. Reference [17] also exemplifies the B-MAC protocol, which performs even better than T-MAC in simulations.

Two other important layers are the *transport layer* and the *network layer*. The transport layer ensures the reliability and quality of data at the source and the sink. Transport layer protocols should have various applications such as packet-loss recovery and congestion control. The network layer consist of routing protocols. It should easily and efficiently propagate data to the base station.

2.2 Localization

Nodes which have been deployed in an ad-hoc manner do not have prior knowledge of their own position. The problem of determining a node's position is referred to as localization. Existing localization techniques include GPS and anchor nodes (e.g. nodes with predetermined positions). The anchor nodes work as reference points for the other nodes who determine their distance from the reference points with message delays.

Another technique is proximity-based localization, which makes use of neighboring anchor nodes to determine their position, and then act as anchor nodes themselves for other neighboring nodes. The GPS and anchor node techniques have their shortcomings, though. The GPS may not work when the nodes are deployed in obstructed areas, and the anchor node technique scales bad in large networks.

One interesting localization algorithm is called *Moores algorithm*. This algorithm, discussed in [7], formulates the localization problem as a two-dimensional graph realization. It does not require anchor nodes, enabling localization without absolute position information.

A graph is, as formulated in mathematical graph theory, a set $G = (V, E)$, where V is an n-dimensional set of vertices, and E is an e-dimensional set of edges. In this case, the vertices represent sensor nodes and the edges represent distances between the nodes. Graph theory provides ways of determining if a given graph has a unique realization and therefore lacks ambiguities. The practical problem is to manage enough distance information by node communication, to make an unambiguous graph of the nodes relative positions in relation to each other.

The algorithm builds the graph by overlapping quadrilaterals. Quadrilaterals is important because they are the smallest possible subgraph that can be unambiguously localized in isolation.

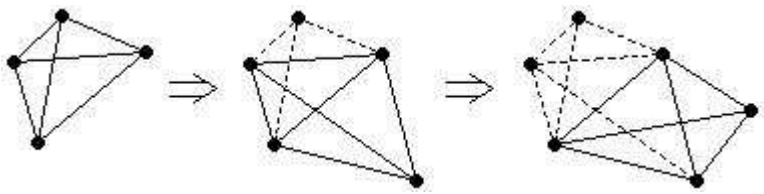

Figure 1: Illustrated localization process with overlapping robust quadrilaterals.

One should distinguish between rigid and non-rigid graphs to understand how to avoid ambiguities. A non-rigid graph can be continuously deformed and hence realize a graph in an infinite number of ways. A rigid graph is only subject to two different types of discontinuous ambiguities; flip ambiguities and flex ambiguities.

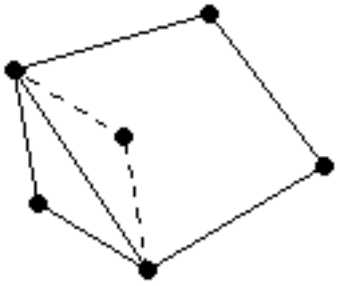

Figure 2: Discontinuous flip ambiguity.

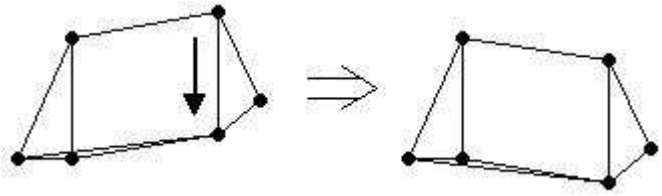

Figure 3: Discontinuous flex ambiguity.

A flip ambiguity is when an individual node also can be realized by its mirror image. A flex ambiguity is when an edge can be removed, making the graph non-rigid, and allowing the edge to be reinserted with the same length at a different configuration. Thus, a graph which cannot become non-rigid by the removal of an edge, formally called a *redundantly rigid* graph, is guaranteed not to have any flex ambiguities.

It is not entirely trivial, but the robust quadrilateral is in fact redundantly rigid, thus excluding the possibilities of flex ambiguities. Flip ambiguities only occur in the algorithm when the distance measurements are noisy. However, [7] shows that one can construct a robustness test where the worst case error probability is bounded when the measurement noise is normally distributed with a known variance.

2.3 Time synchronization

Time synchronization in a WSN is important for power conservation. When a network is synchronized, there will be less collisions and re-transmissions, and therefore, the nodes will be better at cooperating and the communication will go more smoothly. A collision occurs when two nodes transmit at the same time, interfering with each other's transmissions. When that happens, data packets are corrupted, and hence, the energy used during transmission and reception will be

wasted. Routing, which is the process of choosing the paths in which to send network traffic, also relies on time synchronization.

Two ways (among many) to achieve time synchronization is by using Lucarellis algorithm, or Reachback firefly algorithm (RFA), both mentioned in [17]. Lucarellis algorithm is a form of bi-directional synchronization protocol between nearest neighbors, which use timed pulses. If a node sends its pulse out of phase according to a reference pulse from another node, it will increment its pulse phase according to the algorithm.

To get a better grip of how Lucarellis algorithm works, it can be outlined in a more formalized manner. Consider a state variable x_i which increases from 0 to 1, for every node i in the network. When x_i reaches 1, the node emits a pulse and go back to 0. When the nearest neighbor node (say node j) register the pulse signal, its corresponding state variable x_j will make a sudden increase with a small amount ε. If $x_j + \varepsilon > 1$, the node will just reset x_j to 0 and emit a pulse signal. If a small ε is chosen, the time synchronization will be more precise, but slower. In practice, a reasonable approach would be to let ε decrease successively over time in some way.

RFA is an oscillator method inspired by a mathematical model which describes how fireflies and neurons spontaneously synchronize. In some major aspects, RFA resembles Lucarellis algorithm. For a good description of the RFA protocol, see [16]. It is guaranteed that the nodes will converge to synchronicity over time, for both algorithms.

There are several options to resolve the problem of time synchronization used in various synchronization protocols. Reference [11] categorizes several approaches, which include:

- *Master-slave or peer-to-peer synchronization.* In the master-slave approach, one node in the network is master and the rest are slaves. The slave nodes synchronize their clocks with the master node. In practice, nodes with powerful processors and lighter loads are assigned to be masters, since their CPU requirements are higher, generally proportional to the number of slaves. In peer-to-peer, the nodes communicate directly with each other to exchange time information without master-slave relationships, until the whole network is synchronized. The peer-to-peer approach offers more flexibility since it does not risk synchronization prevention due to master node failure. However, peer-to-peer synchronization is more difficult to control.
- *Clock correction or untethered clocks.* Clock correction is when individual nodes adjust their local clock, continually or instantaneously, to keep the network synchronized. Untethered clocks achieve a global time scale without

synchronization, by comparing timestamps between the nodes. The untethered clock approach is becoming popular because of it's energy saving benefits.

- *Internal or external synchronization.* In internal synchronization, one tries to minimize the maximum differences in local time readings between the nodes. External synchronization uses an external time source as a reference, such as the universal coordinated time. The nodes adjust their clocks to this reference time.
- *Probabilistic or deterministic synchronization.* These two methods use a boundary value approach. The probabilistic approach gives an upper boundary for the clock offset with a bounded or determined failure probability, while the deterministic approach guarantees an upper bound for the clock offset with certainty. The reason for using probabilistic methods in WSNs is that they need less data transfer and thus less energy usage.
- *Sender-to-receiver or receiver-to-receiver synchronization.* In the sender-to-receiver method, the sender sends its timestamp to the receiver, which synchronizes its clock to the senders timestamp. To do this, the receiver must consider the message delay time from sender to receiver, which is calculated as the time when the receiver starts requesting a timestamp to when it actually gets one. The problem is that there are variations in message delay time because of network delays and workload in the nodes. This is often solved by calculating the average message delay from many trials. However, in more modern systems, the receiver-to-receiver method is more widely used. This method uses the same broadcast message for many receivers at a time. The message delay is approximately the same for all receivers in single-hop transmission, a fact which is utilized in the method. The receivers then change timestamps with each other and calculates the time offset based on the difference in reception times. This method highly reduces the message delay variance compared to the sender-to-receiver method.

2.4 Compression and aggregation

Data compression is important for WSNs whose batteries are required to last long, since communication is a large energy consumer. By compressing the data before sending it, one can reduce the total energy consumption. The compressing algorithm should not necessarily be optimal in the purpose of compression alone, since the act of compression also consumes energy. The choosing of compression algorithm should be in respect to the criteria of minimizing energy consumption, and therefore, a suboptimal compression algorithm can be preferred. Data aggregation is data from multiple sensors which is combined together and then transmitted. In other words, data aggregation is a form of data compression.

Synopsis diffusion [17] is a framework for aggregation within the network. It consists of three functions: *synopsis generation*, where a synopsis is created from a

sensor reading, *synopsis fusion*, where a new synopsis is created from two others, and *synopsis evaluation*, where a synopsis is translated to its final answer.

The synopsis diffusion process can be divided into two phases: *distribution phase* and *aggregation phase*. During the distribution phase, a query node floods the network with queries (like "average", "sum", "min/max" etc.). Then comes the aggregation phase, where the other nodes use the synopsis fusion function to merge their local synopses with their received synopses. This continues until the query node gets the final synopsis and performs synopsis evaluation on it.

2.5 Automatic control using WSNs

Since WSNs are energy limited, it is preferable to avoid continuous monitoring and only monitor "when necessary", which leads to some drawbacks on control performance. This section is based on [5].

A typical optimal control problem, based on wireless sensor networks, is:

$$\text{Minimize J} = \text{f(Control performance, Energy consumption)} \tag{1}$$

Thus, the criterion gives some tradeoff between control performance and energy consumption. It is desirable that the optimal control problem is formulated in a strict mathematical way. Assume that the process to be controlled is described by the following discrete-time state space model:

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{B}u(k) \tag{2}$$

$$y(k) = \tilde{C}\tilde{x}(k) \tag{3}$$

where $y(k)$ is the control variable, $u(k)$ is the manipulation variable, and $\tilde{x}(k)$ is the state variable. $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ are matrices of suitable dimensions.

To get an analytically formulated optimization problem of type (1), [5] suggests the use of model predictive control (MPC). The idea is to predict future control variables from a prediction model based on future manipulation variables and old control variables, formulate a criterion based on those variables and optimize with respect to $u(k+1), k = 0, 1, \ldots, N$. The optimized $u(k+1)$ is then used for input, and the procedure starts over. The model (2) and (3) is augmented as follows:

$$x(k+1) = Ax(k) + B\Delta u(k) \tag{4}$$

$$y(k) = Cx(k) \tag{5}$$

where $A = \begin{bmatrix} \tilde{A} & \tilde{B} \\ 0 & I \end{bmatrix}$, $B = \begin{bmatrix} \tilde{B} \\ I \end{bmatrix}$, $C = \begin{bmatrix} \tilde{C} & 0 \end{bmatrix}$, $x(k) = \begin{bmatrix} \tilde{x}(k) \\ u(k-1) \end{bmatrix}$ and $\Delta u(k) = u(k) - u(k-1)$. Then the general MPC predictor from step 1 to N is

formulated as:

$$\begin{bmatrix} y(k+1) \\ ... \\ y(k+N) \end{bmatrix} = G \begin{bmatrix} \Delta u(k) \\ ... \\ \Delta u(k+N-1) \end{bmatrix} + Fx(k) \tag{6}$$

where $G = \begin{bmatrix} CB & 0 & ... & 0 \\ CAB & CB & 0 & ...0 \\ ... & ... & & \\ CA^{N-1}B & ... & CAB & CB \end{bmatrix}$ and $F = \begin{bmatrix} CA \\ CA^2 \\ ... \\ CA^N \end{bmatrix}$. Also note that the state variable x is assumed to be perfectly observable.

Let the reference variable be denoted as y_{ref}. A quadratic function J, which should be minimized, is defined as:

$$\mathrm{J} = \sum_{i=1}^{N} \left(y_{ref}(k+i) - y(k+i)\right)^2 + \lambda \sum_{j=1}^{Nu} \Delta u(k+j-1)^2 \tag{7}$$

The quadratic function (7) is minimized with QP (quadratic programming) subject to the following linear constraint conditions:

$$\begin{cases} y_{min}(k+i) \leq y(k+i) \leq y_{max}(k+i) \\ \Delta y_{min}(k+i) \leq \Delta y(k+i) \leq \Delta y_{max}(k+i) \\ u_{min}(k+j-1) \leq u(k+j-1) \leq u_{max}(k+j-1) \\ \Delta u_{min}(k+j-1) \leq \Delta u(k+j-1) \leq \Delta u_{max}(k+j-1) \\ i = 1, ..., N \\ j = 1, ..., Nu \end{cases} \tag{8}$$

The control optimization problem does not have any constraints on energy cost so far. Minimizing (7) subject to (8) can be seen as maximization of control performance, even though a suboptimal control law can be preferred over a wasteful energy consuming optimal control law in a WSN.

To take energy cost (from communication) into consideration, it can be formulated as follows. Define $\mu_C(i)$ as the i-th ahead communication switching variable were $\mu_C(i) = 1$ means communication execution and $\mu_C(i) = 0$ means communication suspension. Also define C_C as some measure on communication cost. The loss function (7) is now reformulated as:

$$\mathrm{J} = \sum_{i=1}^{N} \Big(y_{ref}(k+i) - y(k+i)\Big)^2 + \lambda \sum_{j=1}^{Nu} \mu_C(j) \Delta u(k+j-1)^2 \tag{9}$$

Now there are three conceivable ways to obtain an optimization problem of the form (1), which are listed below:

1. Control performance optimization with communication energy constraint: minimize (9) subject to (8) and $C_C \sum_{i=1}^{Nu} \mu_C(i) \leq C_1$.

2. Communication energy optimization with control performance constraint: minimize $C_C \sum_{i=1}^{Nu} \mu_C(i)$ subject to (8) and $J = \sum_{i=1}^{N} \Big(y_{ref}(k+i) - y(k+i)\Big)^2 + \lambda \sum_{j=1}^{Nu} \mu_C(j) \Delta u(k+j-1)^2 \leq C_2$.

3. Control performance and communication energy optimization: the combined quadratic function $J_C = \sum_{i=1}^{N} \left(y_{ref}(k+i) - y(k+i)\right)^2 + \lambda \sum_{j=1}^{Nu} \mu_C(j) \Delta u(k+j-1)^2 + C_C \sum_{i=1}^{Nu} \mu_C(i)$ is minimized subject to (8).

For every evaluation, an optimal manipulation sequence $\mu_C(1),\ \Delta u(k),\ \mu_C(2),\ \Delta u(k+1), ...,\ \mu_C(Nu),\ \Delta u(k+Nu-1)$ is gained. The first term $\Delta u(k+i-1)$ such that $\mu_C(i) = 1$ is applied after i time steps, and held until the next evaluation is applied.

Reference [5] also briefly discusses the case when the state variable x is not measurable, and must be estimated.

2.6 Sensors in WSNs

It is desirable that sensor technology is provided for continuous sensing of wide varieties of variables, to provide extensive applications of WSNs. So-called passive sensors, which operate without electricity, are very promising for WSN utilization.

Wave technology sensors is a passive sensor type with a wide range of applications, including pressure and torque, temperature, vapor and moisture measurements [2]. The principle of acoustic wave sensors is conversion between mechanical waves and oscillating electrical fields, with the help of piezoelectric materials, such as quartz. When the material in which the wave is propagating is imposed by mechanical stress, the velocity and/or amplitude of the wave changes, and the change in wave characteristics can be used to quantify the stress.

If a coating which absorbs specific biological chemicals in liquids is applied, the pressure on the sensor increases with higher concentrations of biochemicals, and thus, the sensor becomes a biosensor. Of all the known acoustic sensors for concentrations in liquids, the so-called *Love wave sensor* has the highest sensitivity according to [2].

3 Applications in Urban Water Systems

This section lists examples of earlier successful applications of WSNs in urban water systems. Besides from urban water systems, WSNs have had many other successful applications including minimization of energy consumption in Heat Ventilating and Air Conditioning (HVAC) systems [12].

3.1 Leak location in water pipe networks

A common problem for manufacturing municipal water pipe networks is detecting leaks, a result of (natural) pipe deterioration. To detect leaks, acoustic equipment is commonly used such as noise loggers, simple listening devices such as ground microphones, and leak noise correlators. Of the three methods mentioned, leak noise correlators is the most effective and reliable method. The leak noise is measured at two different points, and the signals are sent to the correlator, which determines the position of the leak based on the following expression:

$$p_{leak} = p_1 + \frac{p_2 - p_1}{2} + t_{lag} \cdot v_{soundPipe} \tag{10}$$

where t_{lag} is the time shift of the maximum correlation between the measurement points, p_1 and p_2 are the locations of the measurement points in respect to some reference point, and $v_{soundPipe}$ is the propagation velocity of leak noise.

Since wired leak noise correlation systems are costly and difficult to install, the use of a WSN is highly motivated. The LeakFinder^{RT} system, presented in [3], is such a WSN correlation system with an enhanced correlation algorithm and low frequency vibration sensors. The PipeNet project, presented in [10], is another wireless leak detection system under development. WSNs are not only applicable to leak detection in pipes; they are also promising for flow monitoring, since they can monitor the flow in many locations through a pipe network at a relatively low cost. There are many reasons to conduct flow monitoring, including determination of total system flow, identification of flow capacities through the pipe network, and calibration of flow models [13].

3.2 Reduction of Combined Sewer Outflow (CSO) events

A combined sewer outflow (CSO) event could occur in wet weather, which may result in discharges of untreated water into rivers and other watercourses. The CSO events often occur in combined sewer systems, i.e. the older sewer systems without separation of wastewater and backwater. Only in the United States, 850 billion gallons of discharged untreated water each year from CSO events cause risks of eutrophication, drinking water contamination and human illness.

The combined sewer outflow network (CSONet) in [9] is a WSN for storage control, used to maximize the utilization of the existing storage capacity in the combined sewer systems. In the summer of 2005, a pilot CSONet was deployed in South Bend, IN. Three smart valves were controlled using water level data from sensors within the basin and at the CSO outfall, 3.2 miles away. With this data, the basins upstream could open its smart valve in time, releasing water to the lower basins and hence prepare for the CSO event.

4 Toxic effects on the activated sludge process

This work proposes WSN-based monitoring of substances in wastewater which are toxic to the microbes in the activated sludge process at a wastewater treatment plant. Many pollutants have shown inhibitive effects on the activated sludge process because they are toxic to the microbes. This includes pharmaceuticals, cyanide and heavy metals, among other pollutants.

In [8], the toxicity of cadmium, copper and zinc on the activated sludge process is investigated with artificial wastewater. Concentrations of zinc greater than 3 mg L^{-1} were shown to inhibit the development of microorganisms. Copper inhibited the microorganisms even greater with concentrations of 2 mg L^{-1}. Earlier studies have shown that copper concentrations greater than 63.5 $\mu g\ L^{-1}$ entirely inhibit the development of filamentous bacteria. Some microbiologists believe that filamentous bacteria, due to their thread-like formations, is good for the activated sludge process in moderate amounts, since they provide a skeleton for the flocs and make them more stable against mechanical stress. However, it is well documented that large amounts of filamentous bacteria cause sludge swelling [14], which is a state where the sludge sediments slowly. Cadmium, also with a concentration of 2 mg L^{-1}, inhibited the microorganisms even more. Thus, the following toxicity sequence were concluded: $Cd > Cu > Zn$.

In [6], toxic effects of nickel and copper on nitrifying bacteria are investigated specifically. There was no visible inhibition of nitrifiers until the copper concentration reached 5 mg L^{-1}, or until the nickel concentration reached approximately 100 mg L^{-1}.

A common way to quantify acute toxicity is to estimate the *median lethal dose,* LD_{50}. The LD_{50} variable is defined as the toxic dose required to kill 50% of a specific population. It is usually expressed as mass of toxic substance per mass of test subject.

Accurate threshold values for different toxins are hard to estimate, since side effects from other compounds in wastewater are difficult to predict. A reasonable approach is to test different threshold values for different toxins in individual water treatment plants and see how it works.

5 Methods

In this study, the performance of the proposed storage strategy is investigated with respect to WSN node placements, sensor slowness, sampling times and node energy consumption.

The investigations are executed in a model built in SimuLink. The model can be divided in three sub-models; a sewer pipe net model, a wastewater treatment plant model, and a WSN model. Figure 4 shows a model overview without the WSN.

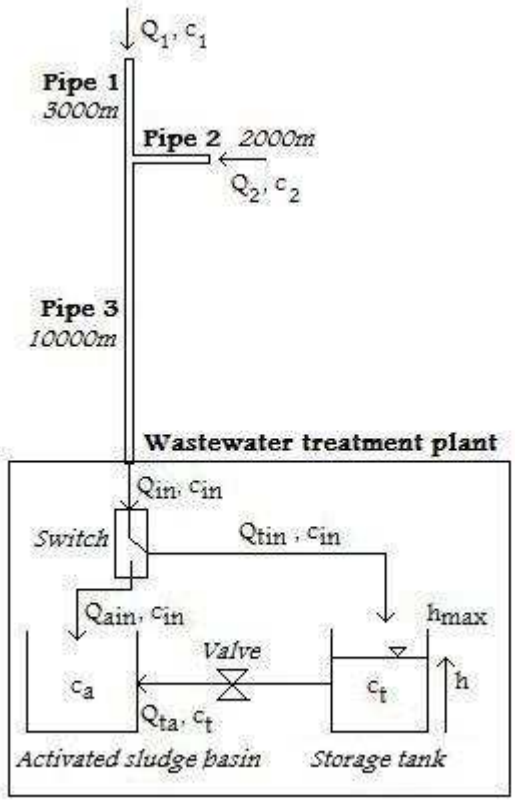

Figure 4: A model overview with model variables. The external model inputs are the flows and toxic concentrations to pipe 1 and 2, referred to as Q_1, c_1, Q_2 and c_2.

The wastewater treatment plant model includes, apart from the activated sludge basin, a storage tank and a small sewer net. The wastewater is bypassed to the storage tank if the WSN registers high toxic concentrations (meaning toxic levels

that violate a certain threshold). The buffer tank empties into the activated sludge basin in small portions to keep toxic levels low.

In the model, only one toxin is present (the choice of toxin is arbitrary). Formations of toxin complexes and other chemical processes in the pipes are neglected; what comes in also comes out. The activated sludge basin is represented by a mass balance equation with a toxic concentration variable, and the challenge is to keep this concentration low. In reality, the proposed system should of course consider several different toxins which could inhibit the activated sludge process, but the conversion of a single toxin-model to a multiple toxin-model is straightforward.

The model builds on the following assumptions:

1. The effluent flow rate from the activated sludge basin is always equal to the total influent flow rate and does not have a maximum limit.
2. The propagation of flow velocity through the pipe is simulated as momentum conversion in each segment. This will be explained further on.
3. The propagation of toxin through the pipe is described by a numeric dispersion algorithm, which depends on the time resolution of the model. This will also be explained further on.
4. The hydraulic dynamics of the pipes within the wastewater treatment plant is neglected. The flow from the inlet to a tank and the flow from the storage tank to the activated sludge tank is hence not delayed due to long pipes.
5. The switch valve and the storage to activated sludge tank valve is simulated as lowpass filters.

5.1 The sewer pipe net model

The purpose of the pipe net model is to provide possibilities of WSN node placements and to get reasonable toxic plume appearances at the wastewater treatment plant inlet. Another purpose is to get a relationship between flow and flow velocity on the form $v = f(Q)$.

The sewer pipe net consists of three pipes put together, as shown in Figure 4. This makes two pipe inlets, a merging point and an outlet to the wastewater treatment plant. The pipes are represented by arrays where each element represent a pipe segment of 100 m. Each segment has a fluid velocity, a fluid cross-section area and a toxic concentration, all of which are varying with time. Note that the flow is the product of velocity and cross-section area.

A set of parameters is needed to be set for the model, such as pipe length, pipe

radius and slope. According to [18], a typical flow velocity is 0.5 m s^{-1}, typical time delays from source to water treatment plant is a couple of hours, and the pipe diameters usually range from 225 mm in the beginning of the net, to 1000 mm at the end of the net. Thus, the pipe radii for pipe 1 and 2 were set to 0.35 m, and the radius for pipe 3 was set to 0.50 m. Note that the radii were chosen comparatively large at the beginning, in comparison to the usual pipe diameters. This is due to the simplicity of the pipe net. In reality, many small pipes merge together into larger pipes, but here, there is only one merging point. If two very small pipes merge into a big pipe, there are not much room for varying flow in the big pipe, since the flow boundaries according to the Manning equation become narrower in the smaller pipes. The pipe lengths were set according to Figure 4, to maintain a time delay of approximately 6 hours.

The flow and velocity relation in the sewers is simulated with Manning's equation (11), which is suitable since the water does not fill the whole pipe area. Manning's equation is given as:

$$v = \frac{1}{n} \cdot R_h^{\frac{2}{3}} \cdot S^{\frac{1}{2}} \tag{11}$$

where v is the mean flow velocity [m s^{-1}], R_h is the *hydraulic radius* [m], S is the slope of the water surface or the *linear hydraulic head loss* [m/m], and n is a dimension free roughness parameter. Reference [1] lists typical values for n for different materials, where 0.016 corresponds to rough asphalt or untreated gunite, and troweled concrete has 0.012, for example. The sewer pipe walls were expected to be made of concrete with a roughness equivalent to rough asphalt due to impacts, roots or other disturbances. Hence $n = 0.016$ was chosen for all pipes.

The slope parameter S was set to 0.001 m/m for pipes 1 and 2, which is shown to give approximately quarter full pipes at flow velocities of 0.5 m s^{-1}. To be more specific, the flow $Q = 3.1818$ m^3 min^{-1} generated the velocity $v = 0.4978$ m s^{-1} and flow cross-section area $A = 0.1065$ m^2 according to the Manning equation, applied as described below. When the flows of pipe 1 and 2 merge into pipe 3, each with flow velocities 0.5 m s^{-1}, the flow in pipe 3 will be approximately 6.4 m^3 min^{-1}. S for pipe 3 is thus chosen so a flow of 6.4 m^3 min^{-1} approximately gives a flow velocity of 0.5 m s^{-1}. Thus, S was set to $6 \cdot 10^{-4}$ for pipe 3. This means that if $Q = 6.7576$ m^3 min^{-1}, then $v = 0.4971$ m s^{-1} and $A = 0.2265$ m^2.

R_h can be further described as $R_h = \frac{A}{P}$, where A is the fluid cross-section area [m^2] and P is the *wetted perimeter* [m], a parameter commonly used in environmental engineering, see Figure 5.

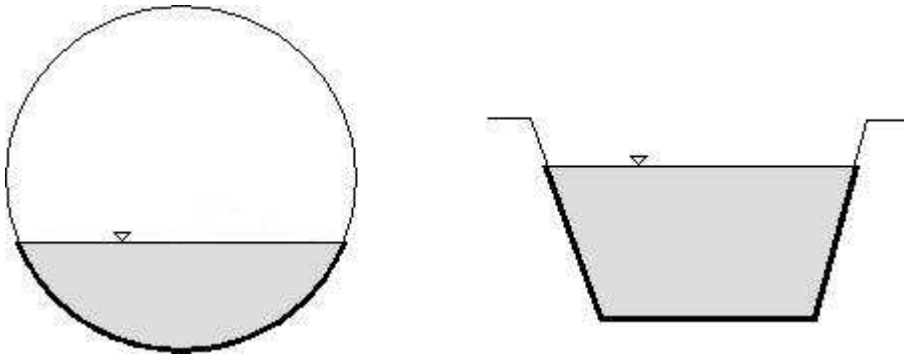

Figure 5: The wetted perimeter.

The wetted perimeter is illustrated in two examples in Figure 5; a pipe to the left and an open canal to the right. Both figures are shown in a view perpendicular to the flow. The wetted perimeter for each case is the length in meters of the fat line, which is the surface line submerged in the fluid.

Since $v = \frac{Q}{A}$, where Q is the flow $[m^3\ s^{-1}]$, (11) can be reformulated as:

$$\frac{Qn}{S^{\frac{1}{2}}} = R_h^{\frac{2}{3}} \cdot A = \frac{A^{\frac{5}{3}}}{P^{\frac{2}{3}}} \tag{12}$$

which will be seen as a more useful formulation if a relation of the form $v = f(Q)$ is wanted. A way to derive A from the above expression (when Q is known) is wanted, with which v easily is derived from $v = \frac{Q}{A}$. The next step is to define the angle θ as illustrated in Figure 6. θ is defined as the angle from the altitudinal line to the merging point between the water surface and the pipe wall, at the center of the pipe. When $\theta = 0$, the pipe is empty, and when $\theta = \pi$, the pipe is full.

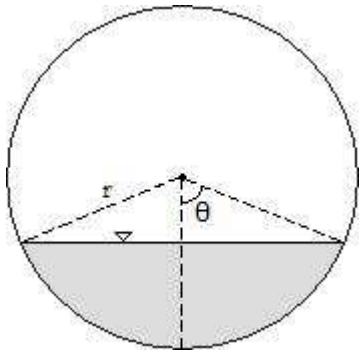

Figure 6: A cross-sectional view of a pipe.

A can unambiguously be expressed with θ as follows:

$$A = r^2(\theta - \sin\theta\cos\theta) \tag{13}$$

which means that the right hand side of (12) expressed as a function of θ would be useful. Since $P = 2\theta r$,

$$R_h^{\frac{2}{3}} \cdot A = \frac{A^{\frac{5}{3}}}{P^{\frac{2}{3}}} = r^{\frac{8}{3}} \cdot \frac{(\theta - \sin\theta\cos\theta)^{\frac{5}{3}}}{(2\theta)^{\frac{2}{3}}} \Leftrightarrow \frac{Qn}{S^{\frac{1}{2}} r^{\frac{8}{3}}} = \frac{(\theta - \sin\theta\cos\theta)^{\frac{5}{3}}}{(2\theta)^{\frac{2}{3}}} = f(\theta) \tag{14}$$

However, it would be more practical to have an explicit expression of θ, since θ is the unknown variable to be derived from Q with the help of (12). But (14) cannot be analytically solved with respect to θ, which means that a numerical solution is necessary. The graph from this numerical solution is shown in Figure 7. This graph contains 100 data points and was generated numerically. The model uses it as a table where θ ranges from 0.1 to π, interpolating between the points.

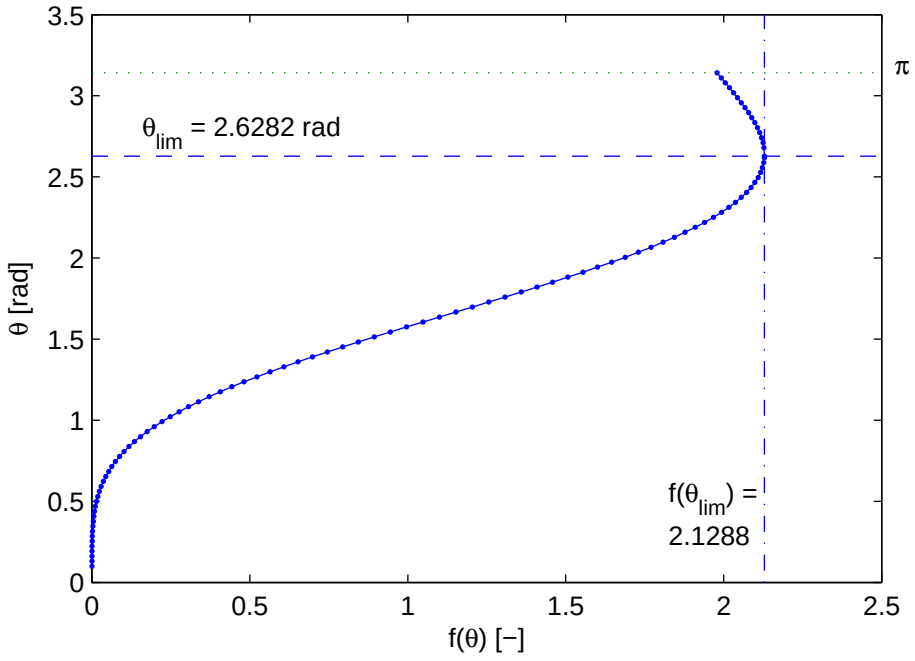

Figure 7: θ vs $f(\theta)$.

The relationship in Figure 7 is ambiguous for some $f(\theta)$. This is because there is a certain θ_{lim} (shown in Figure 7) where $0 < \theta_{lim} < \pi$, with the unambiguous property $f(\theta_{lim}) = \max f(\theta)$. However, θ_{lim} responds to a certain Q_{lim} according to (12): $Q_{lim} = \frac{r^{\frac{8}{3}} S^{\frac{1}{2}}}{n} \cdot f(\theta_{lim})$. If $Q > Q_{lim}$, the upper part of the graph is used, otherwise, the lower part is used. The model does not extrapolate if $f(\theta)$ is outside the boundary; it simply takes the end values instead.

When θ is determined and A is calculated from (13), v is calculated from $v = \frac{Q}{A}$. This is done in the pipe inlets, where Q is an in-parameter to the model. It is also done to calculate initial velocities for all pipe segments, given initial flow values. The propagation of Q and v throughout the pipes is described by a discretized version of the *continuum equation*, and the preservation of momentum principle, respectively. The continuum equation is as follows:

$$\int_{\mathrm{CS}} \rho(\overrightarrow{v} \cdot \mathbf{n}) \, \mathrm{d}A = -\frac{d}{dt} \int_{\mathrm{CV}} \rho \, \mathrm{d}V \tag{15}$$

where ρ is the water density [kg m^{-3}] and $\overrightarrow{v}$ is the relative velocity vector. CS and CV under the integral signs stand for *control surface* and *control volume* respectively. Here, $\mathbf{n}$ is a unity vector perpendicular to an infinitesimal area $\mathrm{d}A$ pointing out from the control volume.

In the model, the control volume is a cylindrical pipe segment of 100 m according to the pipe discretization, whose area is the control area. However, only the areas normal to the pipe flow need to be considered, since there are no flow through the pipe walls (i.e. $\overrightarrow{v} \cdot \mathbf{n} = 0$ since the vectors are perpendicular). Thus, (15) becomes (after elimination of ρ):

$$v_{p-1} A_{p-1} - v_p A_p = \frac{\mathrm{d}V_p}{\mathrm{d}t} \tag{16}$$

where v_{p-1} and v_p are net velocities in and out of the control volume of pipe segment p, and A_p and A_{p-1} are cross-section fluid areas at the control volume in- and outlets. This equation is easily understood intuitively; the positive change of water volume over time equals the net inflow minus the net outflow.

Equation (16) is now discretized with Euler forward as follows, to get a time-update equation for A:

$$\begin{aligned} &\frac{V_p^{t+1} - V_p^t}{t_s} = v_{p-1}^t A_{p-1}^t - v_p^t A_p^t \Leftrightarrow \\ &V_p^{t+1} = V_p^t + t_s(v_{p-1}^t A_{p-1}^t - v_p^t A_p^t) \Leftrightarrow \\ &A_p^{t+1} = A_p^t + \frac{t_s}{l_p} \left(v_{p-1}^t A_{p-1}^t - v_p^t A_p^t\right) \end{aligned} \tag{17}$$

where t_s is the fixed time step, which is set to one minute in the model, and l_p is the pipe segment length, which is set to 100 m as mentioned above. The last

relation comes from substituting $V_p^{t+1} = \frac{A_p^{t+1}}{l_p}$.

The time update equation for v is, as previously mentioned, derived from the momentum conservation principle. The following relation is presumed:

$$\rho V_{1,t} v_t^{p-1} + \rho V_{2,t} v_t^p = \rho V_p^{t+1} v_p^{t+1} \tag{18}$$

where the volumes V_1 and V_2 are explained in Figure 8.

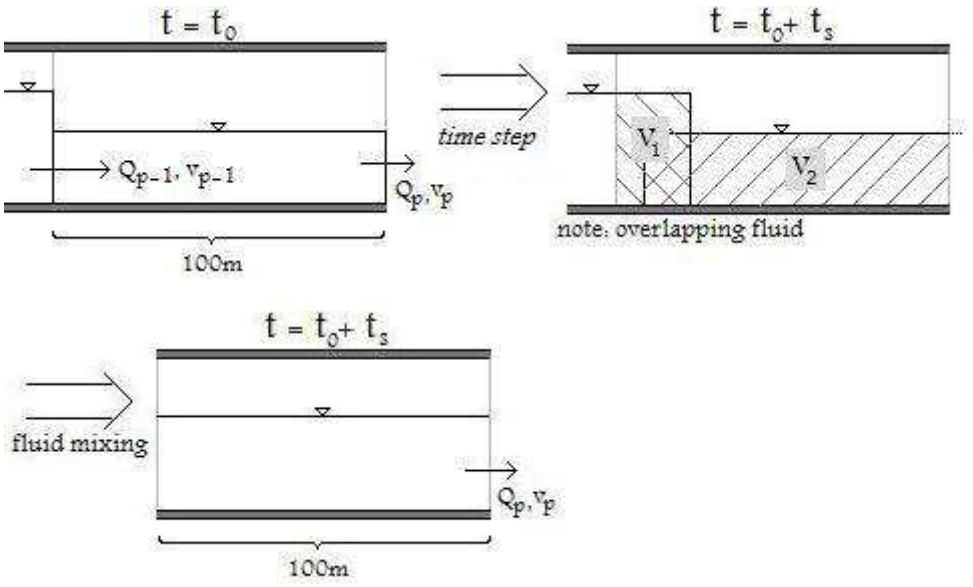

Figure 8: An illustration of the momentum conservation principle, as used in the simulation algorithm.

In Figure 8, there is a time update first, where the fluid blocks move according to the current velocities in the pipe segments. The next step is total mixing, where the total fluid volume in the segment is merged together. This is where the velocity update occurs according to the momentum conservation principle, since the volumes V_1 and V_2 have different velocities. Both the time-update and the mixing step are described by (18). Note that this is a velocity propagation algorithm which is untested empirically. Also note that it depends on the time step used in the model, in other words, the way in which the model is discretized.

When the volumes is merged this way, there is a risk that the merged volume will overlap the total volume possible for the pipe segment, because of overlapping volumes after the time step. This will only happen for big and positive Q gradients, and must be considered when determining the model range for Q.

Substituting $V_{1,t} = A_p^t(l_p - v_p^t t_s)$, $V_{2,t} = A_{p-1}^t v_{p-1}^t t_s$ and $V_p^{t+1} = \frac{A_p^{t+1}}{l_p}$ and eliminating ρ gives:

$$v_p^{t+1} = \frac{A_p^t(l_p - v_p^t t_s)v_p^t + A_{p-1}^t t_s (v_{p-1}^t)^2}{A_p^{t+1}} l_p \tag{19}$$

And then, finally, $Q_p^{t+1} = v_p^{t+1} A_p^{t+1}$.

Another possible selection of a time update equation for v is to use the Manning equation in another way. One could use (17) and calculate θ from $A = r^2(\theta - \cos\theta \sin\theta)$, and then use the Manning equation (11) to get v. However, this means that θ will be calculated from another table, which may result in slight inconsistencies between the two ways the equation is applied. This problem can be avoided by only applying the equation in one of the two ways consequently, and since v is updated from A according to the continuum equation, the first way cannot be used. Because of this, A has to be an input parameter to the model instead of Q as described in Figure 4, and Q will have to be calculated explicitly from $Q = vA$. Secondly, the Manning equation is designed for steady state flow, which means that it is unreliable to use it as a time update equation. Because of this, the momentum conservation principle is used instead, even though it is an idealization since it does not consider friction loss against the pipe walls.

The other input parameter into the pipe inlets, except for Q, is the toxic concentration c [μg L^{-1}]. The propagation of c through the pipe is described as a mass balance equation as follows:

$$\frac{\mathrm{d}(c_p V_p)}{\mathrm{d}t} = \frac{\mathrm{d}C_p}{\mathrm{d}t} V_p + \frac{\mathrm{d}V_p}{\mathrm{d}t} c_p = Q_{p-1} c_{p-1} - Q_p c_p \tag{20}$$

which is discretized by Euler forward:

$$\frac{c_p^{t+1} - c_p^t}{t_s} V_p^t + \frac{V_p^{t+1} - V_p^t}{t_s} c_p^t = Q_{p-1}^t c_{p-1}^t - Q_p^t c_p^t \tag{21}$$

and after some algebraic manipulations, a time update equation for c is given:

$$c_p^{t+1} = \frac{t_s}{A_p^t l_p}(A_{p-1}^t v_{p-1}^t c_{p-1}^t - A_p^t v_p^t c_p^t) + \frac{2c_p^t A_p^t - c_p^t A_p^{t+1}}{A_p^t} \tag{22}$$

This equation depends on the time resolution in an analogous way as the time-update equation for the flow velocity (19). Hence, numerical dispersion, meaning dispersion as an effect of discretization, is used in this model. To achieve dispersion in line with empirical observations, the *advective-dispersive equation* should be used instead.

At the merging point where the outlets of pipe 1 and 2 are the inlet of pipe 3, the following mass balance approaches are used:

$$Q_{in,3} = Q_{out,1} + Q_{out,2} \tag{23}$$

$$c_{in,3} = \frac{Q_{out,1}c_{out,1} + Q_{out,2}c_{out,2}}{Q_{out,1} + Q_{out,2}} \tag{24}$$

5.1.1 Model ranges

The eventual, physical upper range for c would be far over reasonable concentrations. The physical lower range is of course 0. The model ranges of Q come from the fact that the model is not designed to handle full-pipe flows. This also implies ranges for A and v because of the mutual correlations.

The flow inputs to the model, which are the flows at the inlets of pipe 1 and 2, are denoted as Q_1 and Q_2. Consequently, the incoming toxic concentrations are denoted as c_1 and c_2.

For constant flow, the upper limit for Q_1 and Q_2 is 15.6 m^3 min^{-1}, which is Q_{full} for pipes 1 and 2 shown in Figure 9. Due to the risk of overlapping effects, as described in Figure 8, it is not possible to make a table of model ranges for Q_1 and Q_2 for varying flow. Here, Q_1 and Q_2 must be chosen so that the volumes in the pipe segments (for all pipes) does not exceed $\pi \cdot r^2 \cdot l_p$ during the simulations.

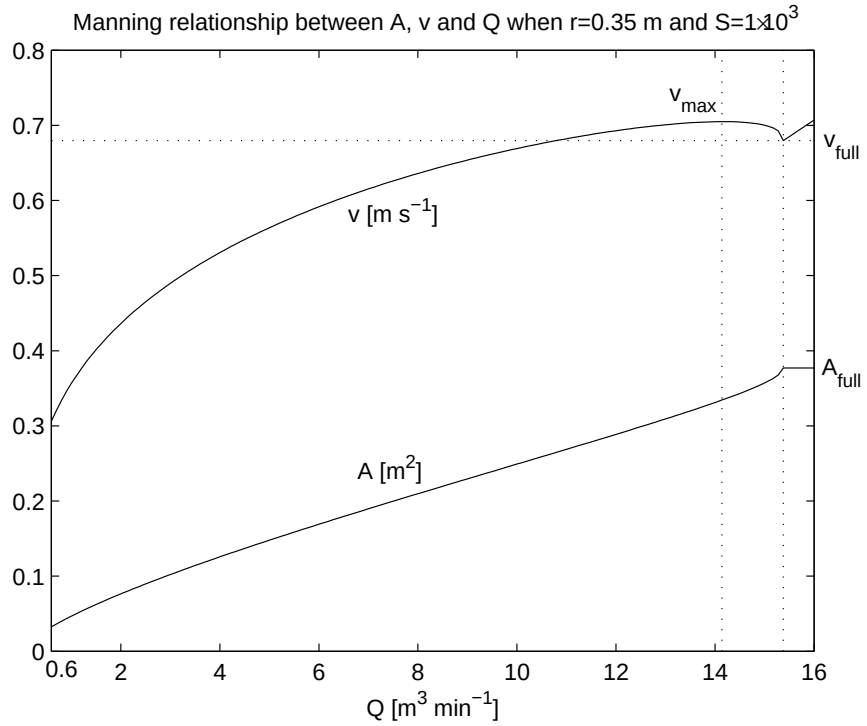

Figure 9: The v, A and Q relationships in pipe 1 and 2.

Figure 9 suggests an upper limit for Q in pipes 1 and 2, since a full pipe should be avoided. Here, $A_{full} = \pi \cdot 0.35^2 \approx 0.38$ m^2, which corresponds to $v_{full} = 0.68$ m s^{-1} $= 40.8$ m min^{-1} and $Q_{full} = 15.6$ m^3 min^{-1}. A natural thing would be to chose Q_{full} as an upper limit for Q throughout the pipe. However, there might be unwanted overlapping effects in the algorithm described by Figure 8, due to high Q gradients in the pipe. As the pipe area fills, there will be a point when the velocity actually decreases due to friction against the pipe walls. This point is called v_{max} in the graph. $v_{max} = 0.70$ m s^{-1} $= 42.0$ m min^{-1}, which corresponds to $Q_{max} = 14.4$ m^3 min^{-1} and $A_{max} = 0.33$ m^2.

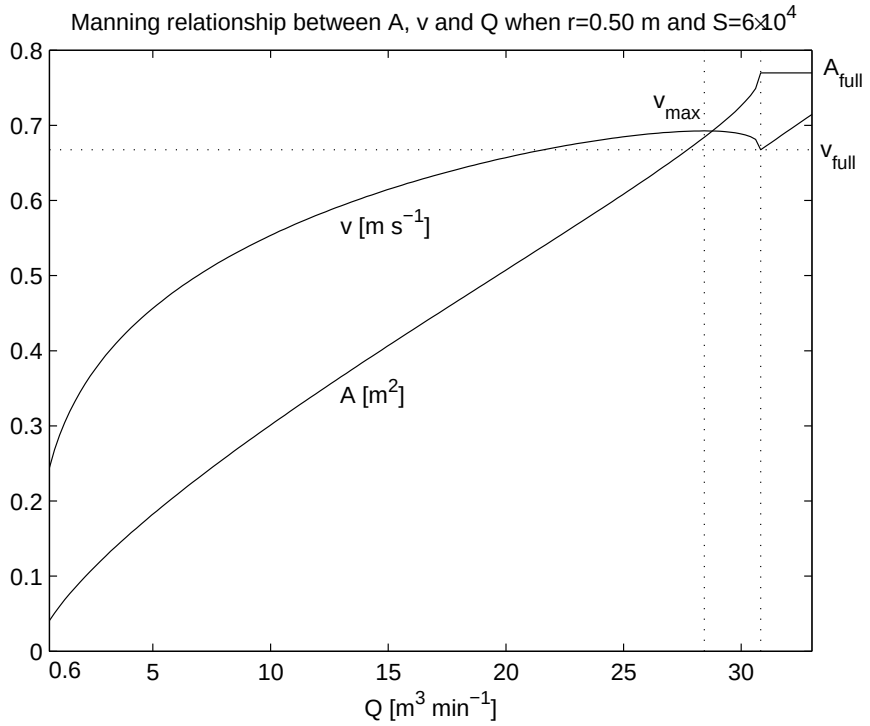

Figure 10: The v, A and Q relationships in pipe 3.

Accordingly, Figure 10 suggests an upper limit for Q in pipe 3. $A_{full} = \pi \cdot 0.50^2 \approx 0.77$ m^2, which corresponds to $v_{full} = 0.67$ m s^{-1} $= 40.2$ m min^{-1} and $Q_{full} = 31.2$ m^3 min^{-1}. Moreover, $v_{max} = 0.69$ m s^{-1} $= 41.4$ m min^{-1}, which corresponds to $Q_{max} = 28.2$ m^3 min^{-1} and $A_{max} = 0.69$ m^2.

5.2 The wastewater treatment plant model

An overview of the wastewater treatment plant model is shown in Figure 4. This part of the model is simulated in variable step time, which means that the incoming variables from the pipe, Q_{in} and c_{in}, are transformed into variable step functions where the current value is held during the pipe model step time t_s.

The activated sludge basin is only simulated as a container with a toxic concentration, c_a. In other words, any microbial activity is not considered in the model. The goal is to keep c_a under a certain threshold, denoted $c_{threshold}$, as much as possible. The dynamics in the basin are described by:

$$\frac{\mathrm{d}V_a c_a}{\mathrm{d}t} = Q_{in,tot}c_{in,tot} - Q_{out,tot}c_a \Leftrightarrow \frac{\mathrm{d}c_a}{\mathrm{d}t} = \frac{Q_{ain}c_{in} + Q_{ta}c_t - (Q_{ain} + Q_{ta})c_a}{V_a} \quad (25)$$

where V_a is the water volume in the activated sludge basin. The *hydraulic retention time*, t_{ret} in the basin is chosen as 2 hours when $Q_{in} = 6$ m^3 min^{-1} ($v_{in} \approx 0.5$

m s^{-1}). Since $t_{ret} = \frac{V_a}{Q_{in}}$, $V_t = 2 \cdot 60 \cdot 6 = 720$ m^3 in the model. Explanations of the other variables in the equation are found in Figure 4. This is in fact a mass balance model with total mixing, where the inflow equals the outflow; hence, V_a is constant. Total mixing implies that the concentration in the outflow equals the concentration in the basin, c_a.

In the storage tank, variable water volume is considered. The water volume V_t equals $A_t h$, where h is the water level and A_t is the storage tank area, set to 600 m^2. The volume of the storage tank is not supposed to limit the performance of the storage strategy in the model. The concentration dynamics in the storage tank are described by the following mass balance equation:

$$\frac{dV_t c_t}{dt} = Q_{tin} c_{bin} - Q_{ta} c_t \tag{26}$$

where the additional variables are explained in Figure 4. Total mixing is assumed also. The product rule gives:

$$\begin{aligned} &\frac{dV_t c_t}{dt} = \frac{dV_t}{dt} c_t + \frac{dc_t}{dt} V_t = \\ &A_t \frac{dh}{dt} c_t + \frac{dc_t}{dt} A_t h = Q_{tin} c_{bin} - Q_{ta} c_t \Leftrightarrow \\ &\frac{dc_t}{dt} = \frac{Q_{tin} c_{bin} - Q_{ta} c_t}{A_t h} - \frac{c_t}{h} \cdot \frac{dh}{dt} \end{aligned} \tag{27}$$

The dynamics of the storage wastewater level h is described by:

$$\frac{dV_t}{dt} = A_t \frac{dh}{dt} = Q_{tin} - Q_{ta} \Leftrightarrow \frac{dh}{dt} = \frac{Q_{tin} - Q_{ta}}{A_t} \tag{28}$$

So, the storage tank dynamics consist of the two differential equations (27) and (28). The variable h should avoid exceeding the tank height h_{max} (which is set to 3 m in the model) but this should be avoided by the control laws. However, both variables h and c_t get a saturation point at zero, so they are constrained to be positive. When $h \to 0$, (27) goes to infinity since h comes in as a denominator. This makes the model very sensible when h is small. To correct for this, the h in (27) is chosen to 0.001 when $h \leq 0.001$.

The dynamics in these tanks are controlled by regulating the switch and/or the storage to activated sludge valve shown in Figure 4. The switch and the valve are both modeled as lowpass filters with time constants 30 and 5 seconds respectively. There is an actuator in the plant which gathers data and sends reference values to the switch and the valve according to some control laws. The reference value for the switch is called ΔQ which tells how the incoming wastewater should be distributed between the tanks. ΔQ is defined by:

$$\Delta Q = \frac{Q_{in} - Q_{ain}}{Q_{in}} = \frac{Q_{tin}}{Q_{in}} \tag{29}$$

The reference value for the valve is simply called $Q_{ta,ref}$ [$\mathrm{m}^3\ \mathrm{min}^{-1}$], where $0 \leq Q_{ta,ref} \leq Q_{max}$. However, since the time constants for the lowpass filters are relatively small compared to the time resolution, there are only small differences between the reference values and the real values.

The control law for ΔQ is given by:

$$\begin{cases} \Delta Q = 0 & \text{if } c_{in,WSN} \leq c_{threshold} \text{ or } h \geq h_{max} \\ \Delta Q = \dfrac{c_a}{c_{threshold} - 1} & \text{if } c_{in,WSN} > c_{threshold} \text{ and } c_a < c_{threshold} \\ \Delta Q = 1 & \text{if } c_a > c_{threshold} \text{ and } h < h_{max} \text{ and } c_{in,WSN} > c_{threshold} \end{cases} \tag{30}$$

where $c_{in,WSN}$ is the incoming toxic concentration measured by the WSN-node at the plant inlet. Additionally, $\Delta Q = 1$ also if $h < h_{max}$ and upstream nodes have registered high c-values some suiting time ago (rules of thumb is used here). The expression $\Delta Q = \frac{c_a}{c_{threshold} - 1}$ makes c_a converge to $c_{threshold} - 1$. The reason for the -1 is to be on the safe side; the storage strategy is not supposed to limit performance in the simulations.

The control law for Q_{ta} is given by:

$$\begin{cases} Q_{ta,ref} = 0 & \text{if } c_a > 0.5 \cdot c_{threshold} + 0.1 \\ Q_{ta,ref} = Q_{ain} \dfrac{0.5 \cdot c_{threshold} - c_{in}}{c_t - 0.5 \cdot c_{threshold}} & \text{if } 0.5 \cdot c_{threshold} \leq c_a \leq 0.5 \cdot c_{threshold} + 0.1 \\ Q_{ta,ref} = Q_{max} & \text{if } c_a < 0.5 \cdot c_{threshold} \end{cases} \tag{31}$$

if $c_a = c_{threshold}$ and the calculated $Q_{ta,ref}$ lands outside the interval $[0\ Q_{max}]$, the model simply takes the correspondent end value of the interval. The expression $Q_{ta,ref} = Q_{ain} \frac{0.5 \cdot c_{threshold} - c_{in}}{c_t - 0.5 \cdot c_{threshold}}$ is used to keep (25) in steady state at the point $c_a = 0.5 \cdot c_{threshold}$. In these conditions, (25) becomes:

$$\begin{aligned} &0 = Q_{ain} c_{in} + Q_{ta,ref} c_t - 0.5(Q_{ain} + Q_{ta,ref}) c_{threshold} \Leftrightarrow \\ &Q_{ta,ref}(c_t - c_{threshold}) = Q_{ain}(0.5 \cdot c_{threshold} - c_{in}) \end{aligned} \tag{32}$$

which gives the control law in (31). The steady state point $c_a = 0.5 \cdot c_{threshold}$ is chosen way below the threshold so there will be no storage emptying before the threshold violation is over.

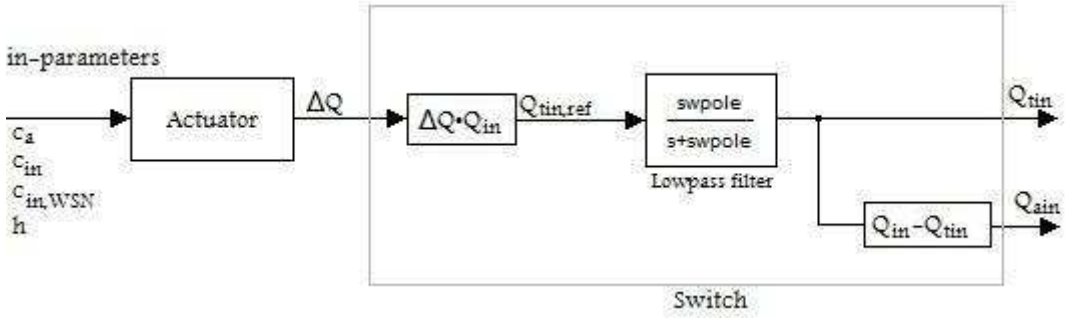

Figure 11: A block diagram of the switch control.

Figure 11 illustrates the switch control law. The control law performed by the actuator is described in (30). The two calculation blocks in the switch can be understood from the relations in (29). The parameter *swpole* equals 2, which corresponds to the time constant $\frac{60}{swpole} = 30$ s.

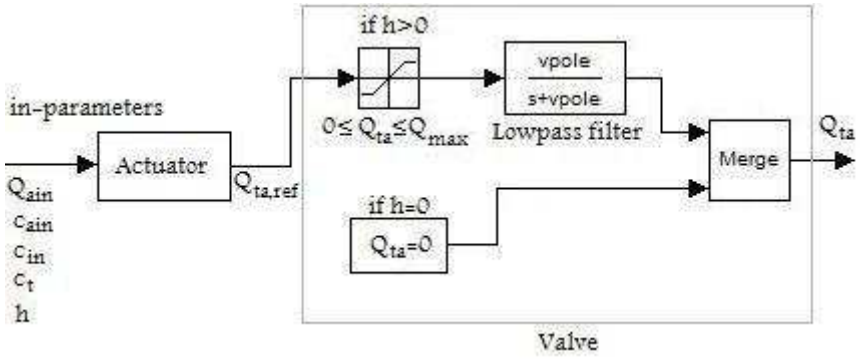

Figure 12: A block diagram of the valve control.

Figure 12 illustrates the valve control law. The control law performed by the actuator is described in (31). The lowpass filter is preceded by a saturation block, keeping Q_{ta} within it's physical limits. If h drops to 0, Qb drops to 0 directly in spite of the lowpass filter. This is done by a merge block in Simulink. The parameter *vpole* equals 12, which corresponds to the time constant $\frac{60}{vpole} = 5$ s.

5.3 The WSN Model

A WSN unit is simulated as a node with a chemical sensor, a processing unit and a radio transmitter/reciever. Sometimes, a flow sensor is also added to the unit.

The nodes can be in active or passive sensing mode, which have a preset *active sample period* and *passive sample period* respectively. Active sample period is sometimes referred to as "Asp" in the text, and passive sample period is sometimes referred to as "Psp".

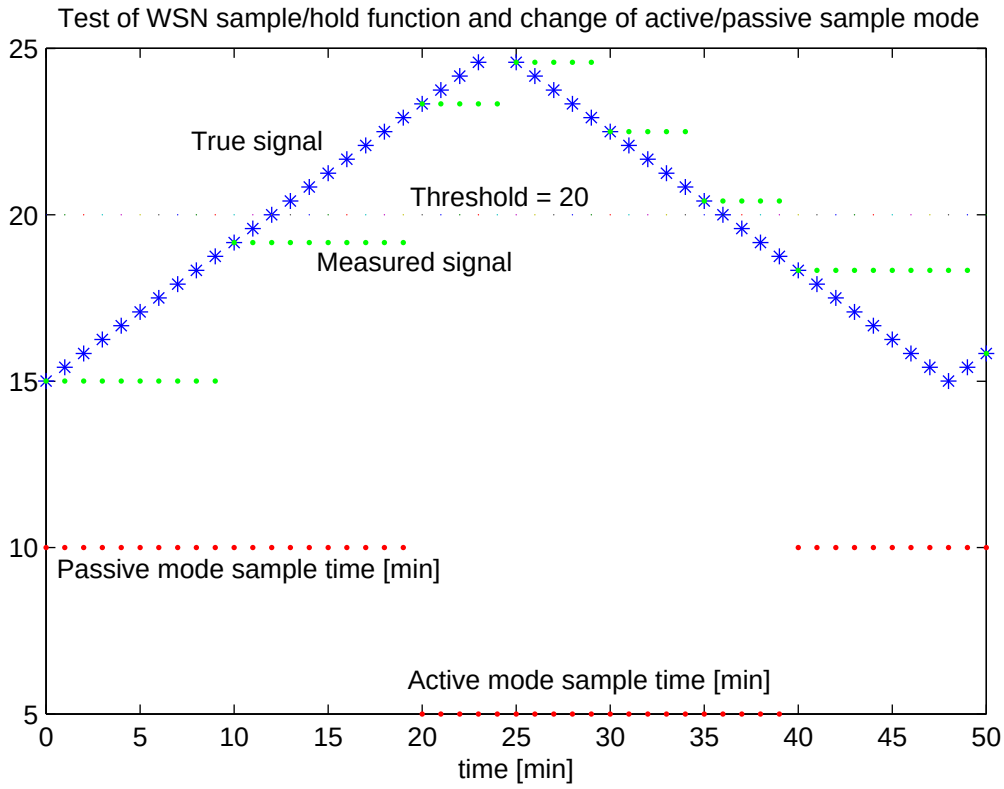

Figure 13: A simulation test of a WSN node sensing a unit-less rise/fall signal.

Figure 13 illustrates the principle of how a single node works. It is applied on a unit-less signal that has nothing to do with the toxic measurements in the pipe model. The node is in active mode if the sensed signal is above the threshold, otherwise it is in passive mode. Here, the active sample period is 5 min and the passive sample period is 10 min. The sensed signal is held during the current sample time.

The qualitative drawbacks for the sensors are simulated by filtering the incoming measurements through a lowpass filter and direct time delays, which represents the slowness of the sensors. It is perhaps a rather ad-hoc way of realizing the drawbacks. The lowpass filter has a rise time of 35.5 minutes for all simulations and the time delay varies from case to case.

The upstream nodes communicate with the downstream nodes by sending alarm signals when registering toxic concentrations higher than the threshold, telling the downstream nodes to go into active mode after some time delay according to some rule of thumb, guaranteeing the downstream nodes to detect the plume also. The alarm signal is in fact a timestamp of the time the toxic plume was detected (when the threshold was violated).

The downstream node will be in active mode until the toxic plume is detected, stay in active mode, and go back to passive mode when the plume passes. It also switches to active mode if it registers threshold violations without getting any alarm signals (which means that the upstream node missed the toxic plume). In real life, one has to apply some strategy if a plume is not detected at all by the downstream nodes (due to dilution and such) for switching back to passive mode, but such a strategy is not considered in this model.

If the downstream node gets the alarm signal, it sends back a notification packet to the upstream node saying that the alarm was received. If the upstream node sent the alarm signal without getting a notification packet, which means that the receiver node had its radio turned off at the moment, it will send another alarm the next minute, keeping the receiver on.

In a hop-by-hop case, where upstream nodes use downstream nodes as part of their transmission path when sending data to the actuator, the upstream nodes will send data repeatedly each minute until they get a notification packet. This is much worse to energy consumption since data packets are much bigger than alarm signals. The downstream nodes only turn their receivers on at the same time as they activate the sensor, which means that the "listening time period" is the same as the sample time period.

To estimate the comparative lifetime between different sensor configurations, an energy consumption model is applied to the WSN model. The lesser the energy consumption, the longer the life expectancy of the WSN. Since only the relative energy consumption between different cases is important, the background energy consumption is set to 0 (which is the energy consumption when the node is doing nothing). The sensors are set to consume 1 mJ when sensing for each variable.

The energy consumption model for the radio transmitter is based on the model used in [4], and it looks like this:

$$E_{transmit} = k \cdot (E_{elec} - e_{amp} \cdot d^2) \tag{33}$$

$$E_{receive} = E_{elec} \tag{34}$$

where $E_{transmit}$ and $E_{receive}$ is the energy used for transmitting and receiving/idle listening respectively. Here, k is the number of bits transmitted, which equals 32

for each data variable or alarm signal transmitted since they are assumed to be 32-bit float numbers, and 1 for notification packets since they are assumed to be boolean.

For each set of data transmitted, a timestamp of 32 bit is also sent. This has no practical meaning in the simulations other than the additional energy cost. For example, if only c is transmitted, the data will be sent in a 64-bit packet. If both Q and c is transmitted, the data will be sent in 96-bit packets.

E_{elec} is the energy consumption in the circuit per data bit, set to 50 nJ bit^{-1} in the model, and e_{amp} is the energy consumption of transmitting per bit and distance squared, which is set to 0.1 nJ bit^{-1} m^2. Finally, d is the transmission distance. E_{elec} and 0.1 nJ bit^{-1} m^2 was chosen the same as they where in [4]. Since received messages only come as boolean, $E_{receive}$ is assumed to be the same regardless of receiving or idle listening events, as shown in (34).

5.4 Scenarios and simulation

To make some suggestions for WSN performance in the current application, some scenarios are simulated where a toxic burst enters the wastewater treatment plant. The simulation time for these scenarios is 900 minutes.

How well the storage strategy performs depends entirely on the quality of the toxic monitoring, since the storage tank is dimensionalized to contain the whole volume of the toxic plume. Initially, the sample time of the WSN nodes is the same as the time resolution for the pipe model, namely one minute. The initial sample time is not supposed to be a limiting factor until it is increased to check changes in performance.

There are no localization problems in the scenarios since the nodes are not spread in an ad-hoc manner. There might be time synchronization problems however, for cases with more than one node.

The following holds for all simulation scenarios, if not specified otherwise. Q_1 and Q_2 are constantly 3 m^3 min^{-1}, and the flows throughout pipe 1 & 2 are also 3 m^3 min^{-1} initially. The initial values for the pipe 3 segments are 6 m^3 min^{-1}, since that is the outcome at the merging point. Note that $Q_{in} = 6$ m^3 min^{-1} throughout the simulations at default. The threshold value $c_{threshold}$ is set to 30 μg L^{-1}. Since there is no measurement noise simulated, one could chose a $c_{threshold}$ close to the background concentration to make the system react fast. However, this would be unrealistic since there is measurement and process noise in reality.

The background toxic concentration is 0.5 μ g L^{-1}, which also is the initial con-

centration in the activated sludge basin and the storage tank for all simulations. The toxic burst comes in after 20 minutes of simulation, as a square impulse in pipe 1, with an amplitude of 10 μg L^{-1} and a duration time of 30 minutes. When the plume reaches the wastewater plant inlet, it will be outspread over a longer time, as shown in Figure 14.

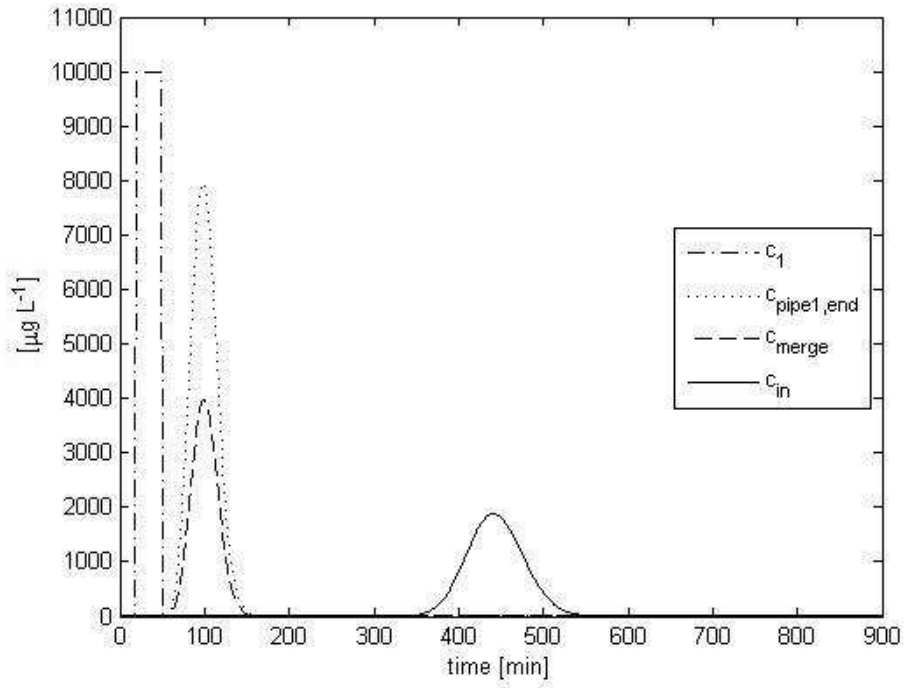

Figure 14: The toxic square pulse in pipe 1 (c_1) and its appearance at the plant inlet (c_{in}).

The pulse is diluted at the merging point, which is shown in Figure 14. This is because of a mixing effect with nontoxic wastewater coming from pipe 2 (c_{merge} is the toxic concentration in the first segment of pipe 3, while $c_{pipe1,end}$ is the end segment of pipe 1 just before the merging point, as the name suggests). The time it takes for the pulse to reach the plant inlet is approximately 350 minutes according to Figure 14.

5.4.1 Performance tests with different sampling times

To estimate the WSN performance on a specific case, the performance needs to be technically defined in some way. Generally speaking, performance in this case is a measure of how much the threshold is violated during the simulation, both the quantity of violation and time duration of violation. A reasonable alternative

would be to choose the area of the c_a-curve that is higher than $c_{threshold}$ in the c_a vs. time graph. That is, the following integral:

$$A_v = \int_{t=0}^{900} \max(c_a(t) - c_{threshold}, 0)\,\mathrm{d}t \tag{35}$$

where A_v is denoted as *area of threshold violation*.

However, it would be even better to chose a parameter that easily can be used for comparison. Therefore, the (negative) measure of performance is furthermore called *threshold violation*, denoted T_{viol}, and defined as:

$$T_{viol} = \frac{\displaystyle\int_{t=0}^{900} \max(c_a(t) - c_{threshold}, 0)\,\mathrm{d}t}{\displaystyle\int_{t=0}^{900} \max(c_{t,no\ temporary\ storage}(t) - c_{threshold}, 0)\,\mathrm{d}t} \cdot 100\% \tag{36}$$

The integral in the denominator is the worst case possible for a given plume; the toxic violation area without any storage strategy at all. For this particular case, $T_{viol} = 100\%$. When there is no threshold violation, which means that the performance is as good as it can get, $T_{viol} = 0\%$. To estimate T_{viol}, the c_a curve must sink back to or below $c_{threshold}$ during the simulation time to get the whole area.

For a given passive sample period, the toxic plume can either be detected early if lucky or late if unlucky. To get a good picture of the performance from different sampling times, simultaneous simulations with different offset times are executed and the T_{viol} from each simulation is gathered. For example, if one wants to test the performance with a passive sample period of 3 minutes, 3 simulations with offset times 0, 1 and 2 minutes are executed. This means that the sensors sample with a three minute interval, where the first sample occurs either initially or after one or two minutes. Since the pipe model has a time resolution of 1 minute, sample times can only be chosen as one minute multiples.

Assume that passive sample period n is chosen. This gives n threshold violation values called $T_{viol}(0)$, $T_{viol}(1), \cdots, T_{viol}(n-1)$ where $T_{viol}(i)$ corresponds to a time offset of i minutes. From this data set, the following parameters are of interest:

$$T_{viol}^{mean} = \frac{1}{n}\sum_{i=0}^{n-1} T_{viol}(i) \tag{37}$$

$$T_{viol}^{var} = \frac{1}{n}\sum_{i=0}^{n-1} (T_{viol}(i) - T_{viol}^{mean})^2 \tag{38}$$

$$T_{viol}^{max} = \max_{0\leq i\leq n-1} T_{viol}(i) \tag{39}$$

$$T_{viol}^{min} = \min_{0\leq i\leq n-1} T_{viol}(i) \tag{40}$$

where "var" in (38) stands for sampled variance.

$T_{viol}^{min} = 0$ for every performance test executed in this report, since they are executed only in cases where the threshold is not violated when sample time is one minute. When there is only one node, active and passive sample periods are chosen equal since there is no point in doing it another way. The only crucial thing is the detection of the plume, which depends on the passive sample period alone.

There is also need for another definition, namely the *sample period range*. In the following text, the term sample period range means the interval of sample times for which

$$T_{viol}^{mean} = 0 \tag{41}$$

$$T_{viol}^{var} = 0 \tag{42}$$

$$T_{viol}^{max} = 0 \tag{43}$$

$$T_{viol}^{min} = 0 \tag{44}$$

The upper limit of this interval, for which (41)-(44) holds, is in the following text referred to as the *sampling period limit*, or $t_{s,lim}$. Correspondingly, the following parameters are also calculated during simulation:

$$E_{tot}^{mean} = \frac{1}{n}\sum_{i=0}^{n-1} E_{tot}(i) \tag{45}$$

$$E_{tot}^{var} = \frac{1}{n}\sum_{i=0}^{n-1} (E_{tot}(i) - E_{tot}^{mean})^2 \tag{46}$$

$$E_{tot}^{max} = \max_{0\leq i\leq n-1} E_{tot}(i) \tag{47}$$

$$E_{tot}^{min} = \min_{0\leq i\leq n-1} E_{tot}(i) \tag{48}$$

where E_{tot} stands for the sum of the node energy consumption during 900 minutes of simulation. This is, of course, only the energy consumption from sensing, idle

listening, receiving and transmitting. $E_{tot} = 0$ only means that the node neither sense, listen, receive or transmit during the whole simulation period, it does not say anything about energy consumption from other activities (such as driving the internal clock, being in sleep mode, running synchronization algorithms, processing data and such). This parameter is only of interest for comparison.

5.4.2 Realization of time synchronization

In the simulations, there will also be energy consumption tests for badly synchronized WSNs, consisting of two nodes which is the upstream node and the plant inlet node. Strictly speaking, the nodes are considered synchronized when they have the same offset times (and, of course, the same passive sample periods). For example, if both nodes have $Psp = 10$ minutes and a time offset of 4 minutes, they will sample at simulation times 4, 14, 24 and so on.

The unsynchronized WSNs are realized in two different ways. In the first approach, the probability that one specific node is ahead of the other is assumed to be equal. If $E_{tot,in}(i,j)$ and $E_{tot,up}(i,j)$ is defined as the total energy consumption during 900 minutes for the plant inlet and the upstream node correspondingly, with inlet node time offset $= i$ minutes and upstream node time offset $= j$ minutes, and if the time offset of the upstream node is at max x minutes higher or lower than the time offset of the plant inlet node, the following is applied:

$$E_{tot,in}(i) = \frac{1}{2x+1} \sum_{j=i-x}^{i+x} E_{tot,in}(i,j) \tag{49}$$

$$E_{tot,up}(i) = \frac{1}{2x+1} \sum_{j=i-x}^{i+x} E_{tot,up}(i,j) \tag{50}$$

(45)-(48) is then used to get the parameters of interest for both nodes.

In the second approach, the upstream node is assumed to be the master node and thus ahead of the plant inlet node (the slave node). If the upstream node is x minutes behind, $E_{tot,in}(i, i-x)$ and $E_{tot,up}(i, i-x)$ are gathered and applied in (45)-(48). This is perhaps more realistic if some time synchronization algorithm is used.

5.4.3 Flow measurements with upstream node compensation

In the third scenario, which is described further on, flow sensors are added to the two nodes, to handle an environment of unreliable flow velocities. These flow measurements are used to estimate delay times used to say how long the inlet node or the plant actuator should wait before reacting to an alarm signal (timestamp) from the upstream node. Figure 20 in Example scenario 2 describes how these

time delays are estimated.

The inlet node will be too slow to handle the incoming flow pulse. To solve this problem, the upstream node will send its data to the plant actuator, which switches to the storage tank after some time delay, independent of inlet node measurements. When the upstream node does this, it will not send any alarm signals to the inlet node.

5.4.4 Example scenario 1: Constant flow with one sensor

In this simulation, the chemical sensor at the plant inlet (which is the only sensor in the scenario) has a time delay of 20 minutes. Recall that the measurements are both delayed and filtered by a lowpass filter with rise time = 35.5 minutes. The delay and lowpass filter represents the quality drawbacks of the sensor. A time delay of 20 minutes is necessary to get threshold violations for WSN measurements. Otherwise, the WSN will react before c_a has reached the threshold.

First, a simulation without any storage strategy was conducted. Figure 15 shows the results. Have in mind that the rate of toxic concentration in the tank is a function of the hydraulic retention time, which is chosen as 2 hours.

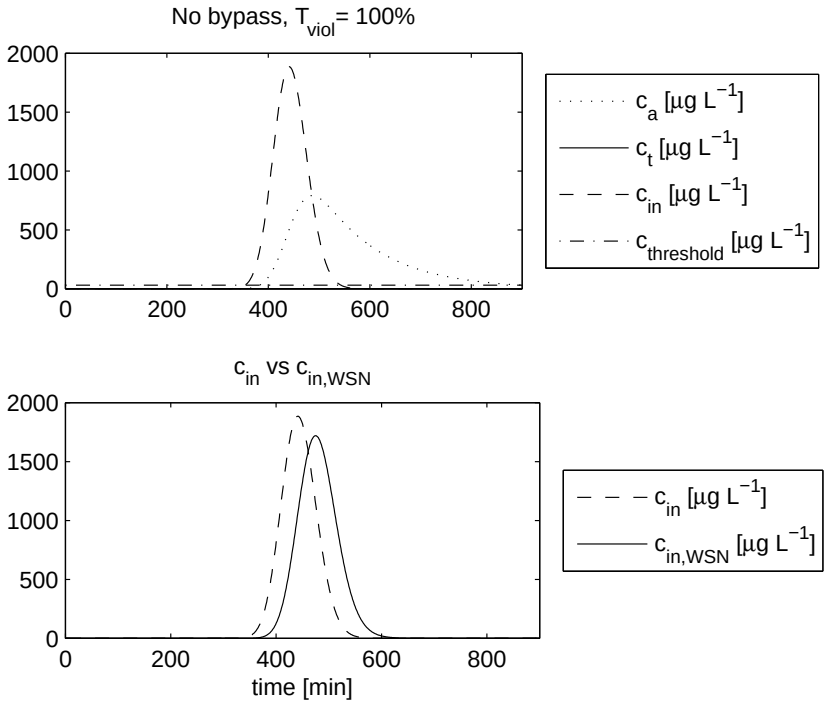

Figure 15: System response without storage strategy.

The upper graph in Figure 15 shows the toxic concentration in the tank when there is no storage strategy at all. T_{viol} mentioned above the graph is the threshold violation parameter defined in (36). The lower graph shows the true c_{in} and the measured $c_{in,WSN}$.

Next, the plant was assumed to have access to ideal measurements at the plant inlet. This means that the measured toxic concentrations equal the true toxic concentrations. The results are shown in Figure 16.

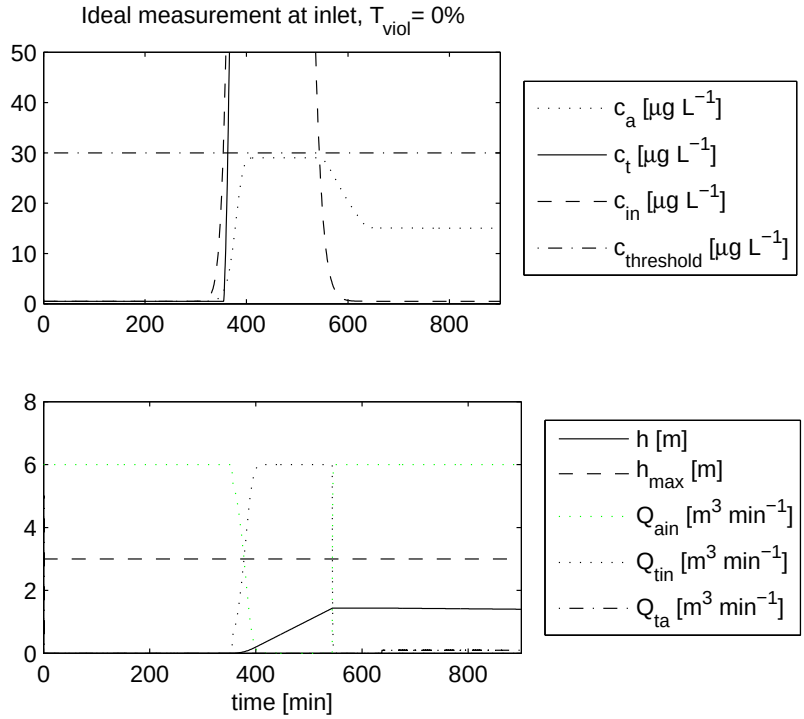

Figure 16: System response with ideal measurements.

The upper graph in Figure 16 shows the lower part of the c_{in} burst shown in Figure 15. The system reacts at the time when c_t is rising rapidly, which is the same time when the storage water level h in the lower graph starts to rise. Here, c_a does not violate the threshold, which is to be expected since ideal measurement is assumed (meaning that the true c_{in} equals the measured c_{in}).

Next, the system had access to WSN measurements at the inlet, instead of ideal measurements. The results are shown in Figure 17.

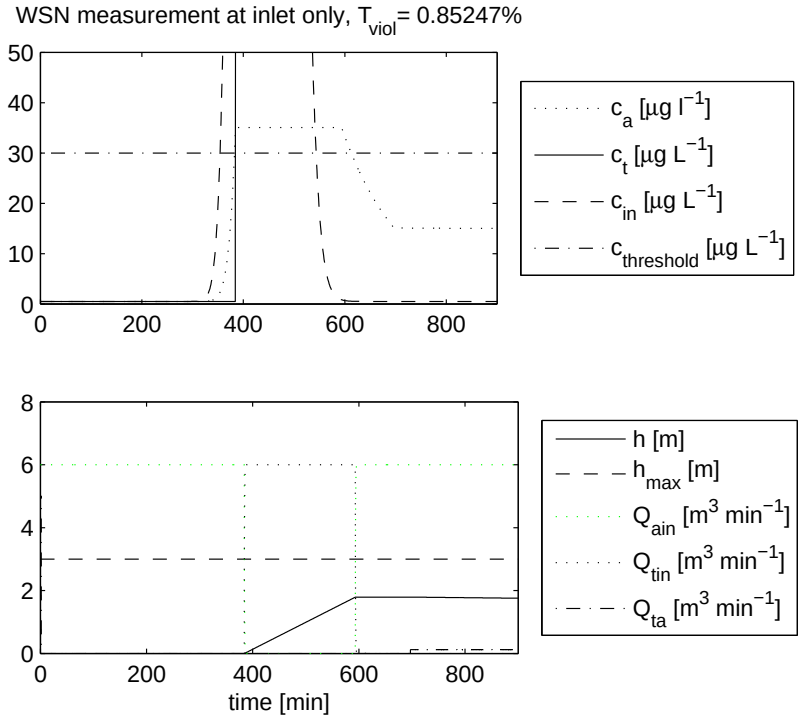

Figure 17: System response with WSN measurements at the inlet.

There is a temporary violation of the threshold due to the slow reaction, as shown in Figure 17. This is corrected later due to dilution from the nontoxic wastewater, however, this dilution takes place a little late.

With sensors like these, it is not a good idea to place them at the inlet. A better idea may be to place them a bit upstream in the pipe, so they have time to react. Since the sensor has a time delay of 20 minutes and the water velocity is approximately 30 m min^{-1}, a reasonable suggestion is to put the node $20 \cdot 30 = 600$ m upstream. This gives the following results, shown in Figure 18:

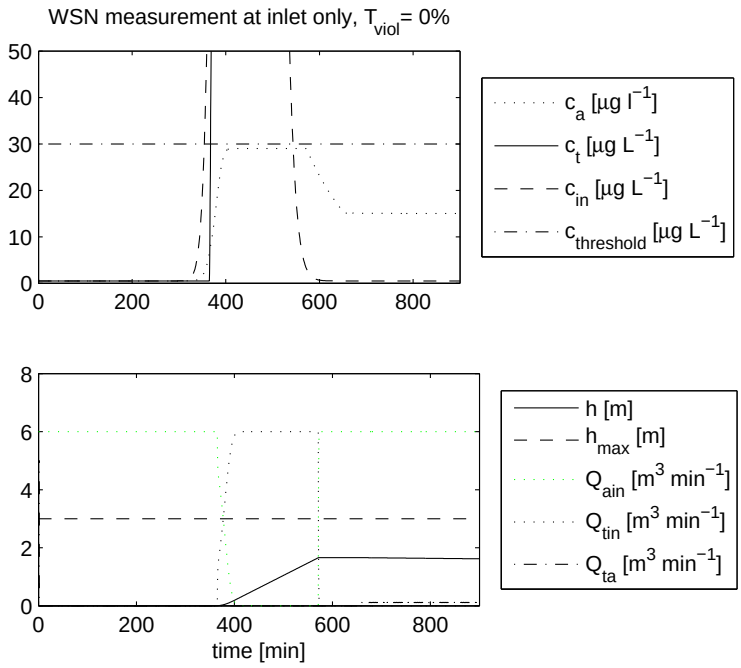

Figure 18: System response with WSN measurements 600 m upstream.

In Figure 18, the threshold is not violated. The system reacts a little slower than it does for the ideal case, but it still has good margins.

The sample period range, defined by (41)-(44), for the case in Figure 18 is the interval $1-17$. This means that the sampling period limit is $= 17$ minutes for this case. If sample time $= 18$ minutes, $T_{viol}^{mean} = 9.9466 \cdot 10^{-4}$ %, $T_{viol}^{std} = 0.0042$ % and $T_{viol}^{max} = 0.0179$ %.

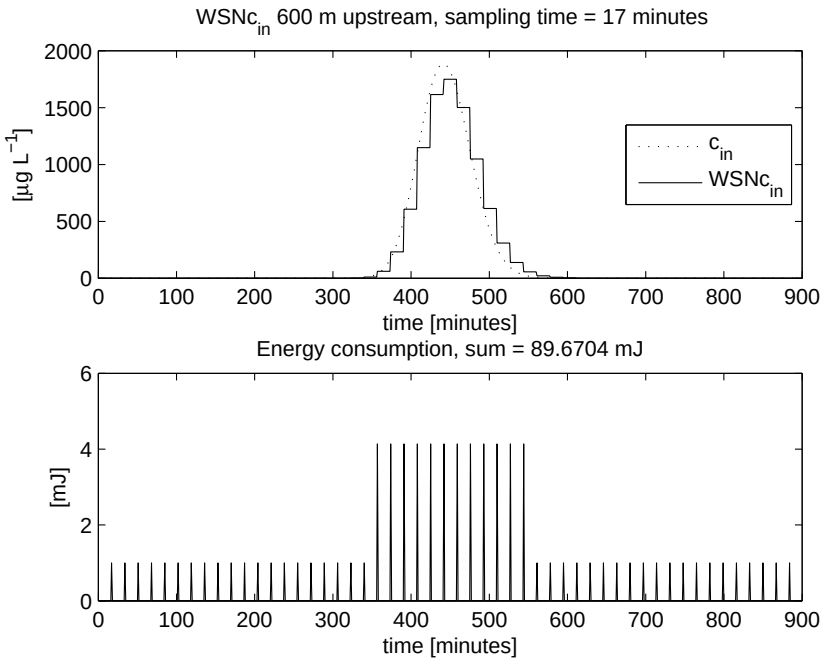

Figure 19: Energy consumption graph with a sample time of 17 minutes.

The upper graph in Figure 19 shows the appearance of the sensor signal, and the lower graph shows the energy consumption of the node, when the sample time is 17 minutes and the time offset is 0 minutes. There are exactly 17 minutes between the energy consumption peaks in the lower graph, which first come from sensor measurements only. The higher peaks come from both sensing and transmission of data, which occur when high toxic concentrations are registered.

Figure 19 is just for illustration; the parameter considered consequently throughout the simulations is the sum of the total energy consumption for 900 minutes, from here on referred to as E_{tot}.

5.4.5 Example scenario 2: Constant flow with two sensors

There exist two main reasons to add more than one node to the pipe network:

- To increase redundancy. If one node malfunctions, the others can take its functional place. This feature is not considered in this work.
- To reduce node energy consumption (on an individual basis), for example by having higher sample time periods (lower sample frequencies) than reasonable for a one-node configuration. Energy consumption reduction implies increase of system lifetime.

When two sensors are employed, a wide variety of WSN strategies can be applied. The storage strategy can take data from upstream nodes into consideration (after some time delay), which requires the upstream nodes to send data. In reality, the data could be sent to the actuator which performs model calculations to predict the concentration at the plant inlet.

If the flow velocity is not constant, there might be a need to measure that variable also, which adds to the system energy consumption. If data is sent by the upstream node, it can either be sent directly to the plant actuator, or to the inlet node in a hop-by-hop strategy. This decreases the energy consumption of the upstream node since the transmission distance decreases, at the cost of higher inlet node energy consumption.

In this model, hop-by-hop routing is not considered since the inlet node is just 100 m from the actuator, which means that additional gains in energy saving are assumed to be small. If the storage strategy only considers inlet node data, the function of the upstream node is only to send alarm signals to the inlet node, making it register the toxic pulse sooner.

The nodes have a time delay of 15 minutes in this scenario. The further upstream the upstream node is placed, the more energy consumption due to long distance transmission will be needed, and the more undeformed the plume will be. In this example, the upstream node is placed 10 000 m from the plant actuator, which is 9900 m from the plant inlet node.

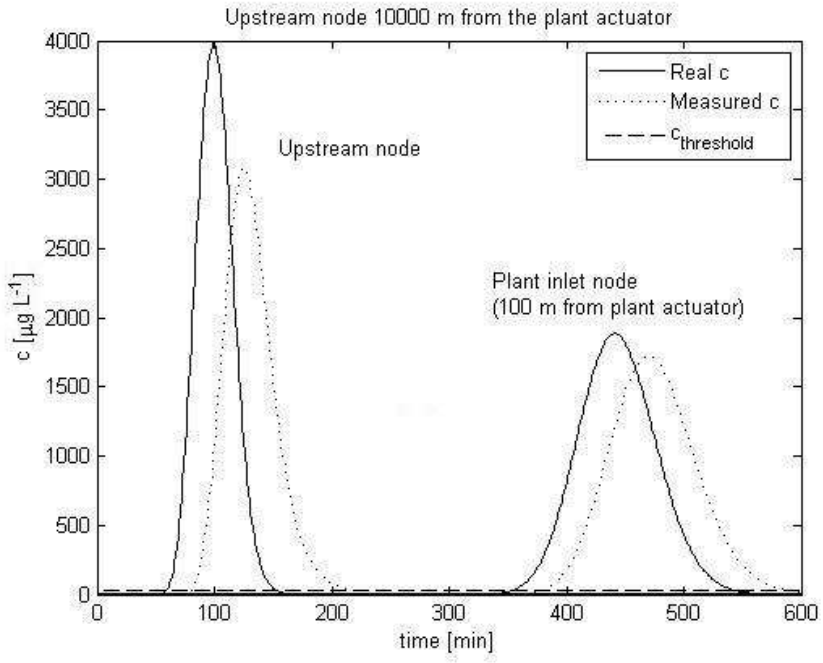

Figure 20: Toxic plume appearances at node locations.

Figure 20 shows the same plume at two different times and locations; one at the first section of pipe 3 just below the merging point where the upstream node is, and the other at the last section of pipe 3 where the plant inlet node is. Both nodes have a sample period of 1 minute. When the upstream node detects the plume (which is when the measured curve first violates the threshold), it takes exactly 299 minutes for the plume to be detected by the plant inlet node downstream. This would suggest the rule of thumb that the inlet node switches to active mode 299 minutes after the upstream node detects the plume.

However, if the passive sample period (Psp) is more than 1 minute, let say 5 minutes, then there is a time of detection interval from 295 to 299 minutes, depending on the sampling time offset. This means that the node should go into active mode 295 minutes after the timestamp value, to be on the safe side.

The time delay from an alarm signal (time stamp) to when the node switches to active mode is furthermore referred to as *response time delay*, or t_r.
To illustrate the consequences of unsynchronized systems, with long sample periods, a simulation was conducted. The results are shown in Figure 21. The choosing of Psp, Asp, time offsets and response time delays for this simulation are discussed below.

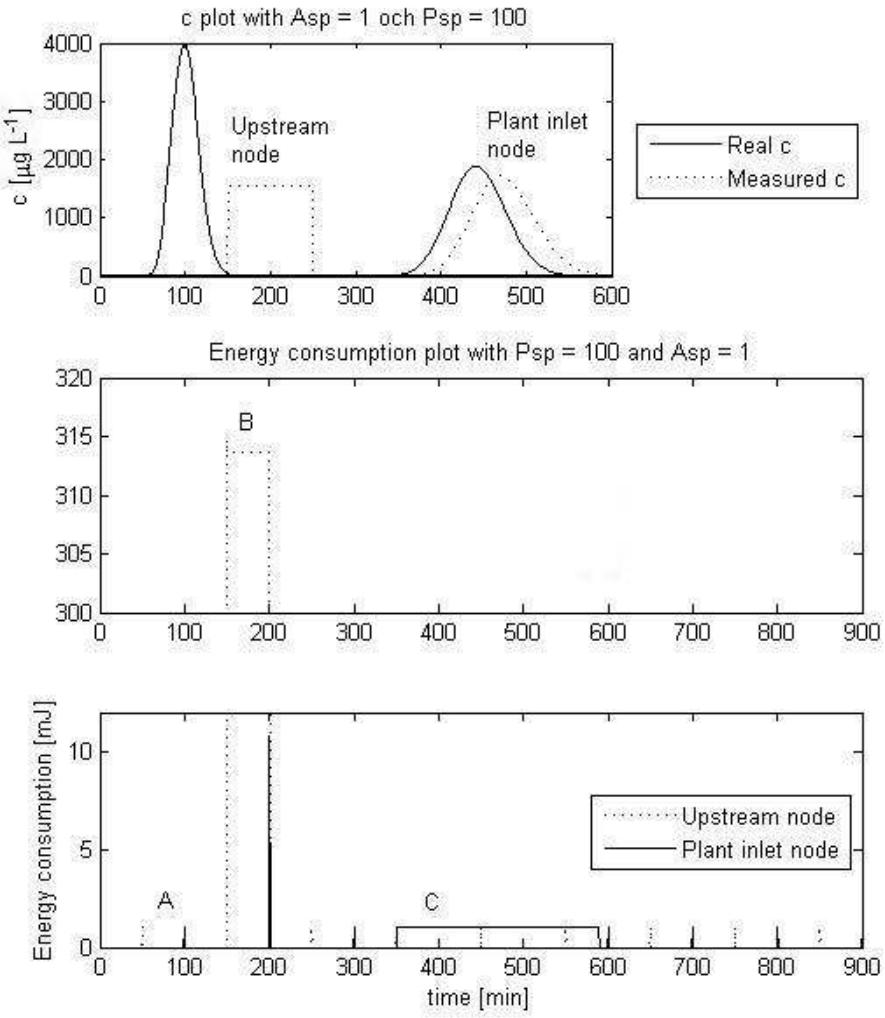

Figure 21: Energy consumption plot.

The upper plot in Figure 21 shows measured vs. real c when $Psp = 100$ and $Asp = 1$. The upstream node detects the plume 150 minutes into the simulation. That measurement is held until the next measurement 100 minutes later. The other plots show the energy consumption of both nodes.

"A" in the lower graph in Figure 21 shows the energy consumption of measurements taking place. The inlet node also turns on its radio during this minute, so its peak is actually slightly higher. The sample time of the upstream node is offset by 50 minutes.

"B" shows the energy consumption peak of the upstream node detecting the plume, turning its radio on (to listen for notification packets) and sending an alarm signal (as a timestamp, which is the detection time = 150 minutes) to the inlet node. Since there is a time offset of 50 minutes, the upstream node tries to send the timestamp 50 times until it is received by the inlet node. The solid peak at the foot of the much larger dotted peak shows the energy consumption of the inlet node measuring, turning its radio on and sending a notification packet.

After some time delay, the inlet node switch to active mode, which is shown in "C". This also explains why the inlet measurement is so smooth in the upper graph. The response time delay (t_r) is chosen as 300 minutes minus Psp and the difference between timestamp and receival time (see Figure 20). Since $Psp = 100$ minutes, the timestamp = 150 minutes and receival time = 200 minutes, the node will wait $300 - 100 - (200 - 150) = 150$ minutes before switching to active mode, from the time when the alarm signal was received.

The additional energy to send the data to the actuator, which starts after 380 minutes when $c_{in,WSN}$ violates the threshold, is not seen in Figure 21 since it is so small. This is because the actuator is only 100 m away from the node. In this simulation, E_{tot}(upstream node) = 16.0 J and E_{tot}(inlet node) = 270 mJ.

The estimated delay time from when the upstream node registers the plume, to when the inlet node registers the plume, is furthermore denoted $t_{est,D}$, where 'D' stands for detection. Additionally, the estimated delay time from when the upstream node detects the plume to when the node reaches the plant inlet (which is sooner than the inlet node actually detects it) is denoted $t_{est,in}$.

The time delay estimation $t_{est,in}$ is used by the actuator in Scenario 3 (which is described further on), when it gets measurements from the upstream node. It is used in an analogous way as how the inlet node handles alarm signals, except that it is a time delay from when the upstream node detects the plume, to when the incoming water switches over to the storage tank independently of the inlet node measurements.

Both $t_{est,D}$ and $t_{est,in}$ are determined as the time delays from upstream node toxic detection to downstream node toxic detection or ideal measurement detection at the inlet, respectively (when $Psp = Asp = 1$ minute). These are the time delays found in Figure 21 for this case. The response time delay is chosen as $t_r = t_{est,D} - Psp$ minutes throughout the simulations.
In reality, $t_{est,D}$ and $t_{est,in}$ is a function of flow velocity, node distances, sensor slowness and sample period. In scenario 3, the flow measurements are assumed to be used for simulation by the actuator to estimate $t_{est,in}$. In reality, larger sample periods are likely to make these estimations worse, but here, the available measurements are assumed to be enough.

5.4.6 Implementation of main scenarios

It could be convenient to have one sensor at the wastewater treatment plant inlet, since toxic conditions can change during the flow path in reality due to chemical reactions or toxic injections close to the plant. Therefore, the first scenario is a scenario with a single node at the plant inlet. This node senses and transmits toxic concentrations only, together with timestamps. The toxic concentration variable together with the timestamp are assumed to consist of two 32-bit floats.

Toxic violations and total energy consumptions of the node during the simulation time are gathered for different sample times and different sensor delay times. This is done for both default Q_{in} and doubled Q_{in} (which is $Q_{in} = 12$ m^3 min^{-1}). The active sample periods and passive sample periods are set equal. This is because the node will go to active mode only when the toxic plume already is detected. However, for a case when T_{viol} exceed 0 %, one can theoretically improve performance by decreasing the active sample period. This is because the sensor could miss the point there c_{in} decreases below $c_{threshold}$. Some cases of this sort are also investigated by decreasing the active sample period. These simulations are executed for sensor time delays of 5, 10 and 15 minutes respectively.

Scenario 2 is a case with two nodes, one at the inlet and the other at varying locations upstream. The simulations are executed with a sensor time delay of 15 minutes. Both sensors have the same time delays and, since time synchronization is assumed, the same time offsets. The supposed increase of system longevity due to energy consumption reductions in individual nodes is investigated. Here, the upstream node sends no data to the plant actuator, its only function is to send alarm signals to the downstream node when it detects toxic levels that violate the threshold.

The active sample period is always chosen as the corresponding $t_{s,lim}$ of the inlet

node. Supposed increases in energy consumption, when the WSN is unsynchronized, are also investigated according to section 5.4.2. Here, the upstream node only measures c.

Scenario 3 is the same as scenario 2, except that the toxic plume comes together with a flow pulse temporarily doubling the flow and making the thumb rules for the estimation of plume arrival at the plant inlet more unreliable. This is due to the fact that the nodes lack flow sensors and thus assume default flow velocity throughout the pipes in this scenario. This is compared with a case when both nodes also have flow sensors, and thus increasing the sensor and transmission energy consumption since there is one more variable to handle (there is one timestamp for each pair of variables).

6 Results and discussion

First, a performance test was conducted. Psp increased from 1 and up, and mean values of T_{viol} and E_{tot} were gathered (see section 5.4.1 for details). For each simulation, $Asp = Psp$. This was done with sensor time delays of 5, 10 and 15 minutes each. The results are shown in Figure 22.

6.1 Scenario 1: One node at the plant inlet

Figure 22 shows a tradeoff relationship between performance and energy consumption (or system lifetime). The threshold violation limits ($t_{s,lim}$) are the sample times of the points coinciding with the x-axis, however, no sample times are shown in this graph but in Figure 23. Each point differs from the neighbors with a sample time of 1 minute.

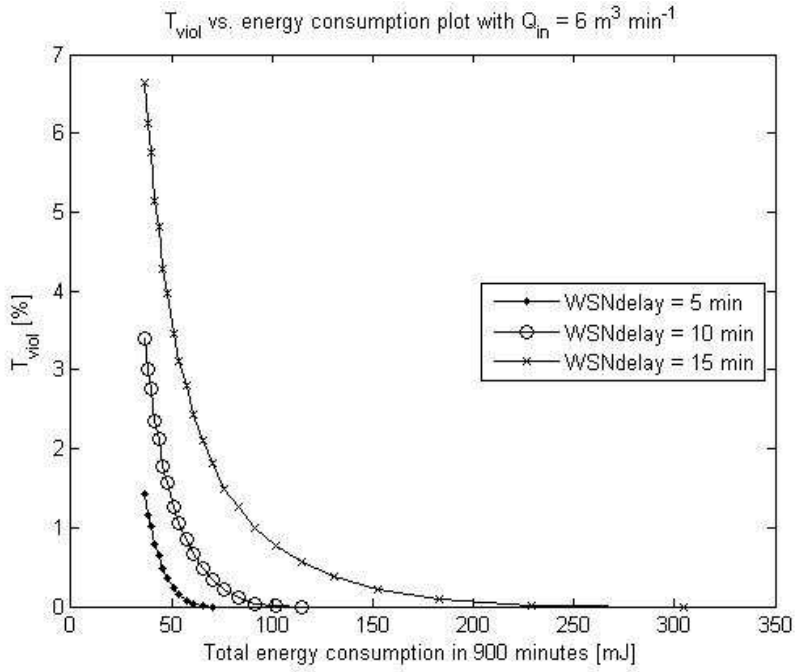

Figure 22: Mean T_{viol} vs. energy consumption.

The same performance test was conducted with double inflow ($Q_{in} = 12$ m min^{-1}). The resulting changes in threshold violations are shown in Figure 23.

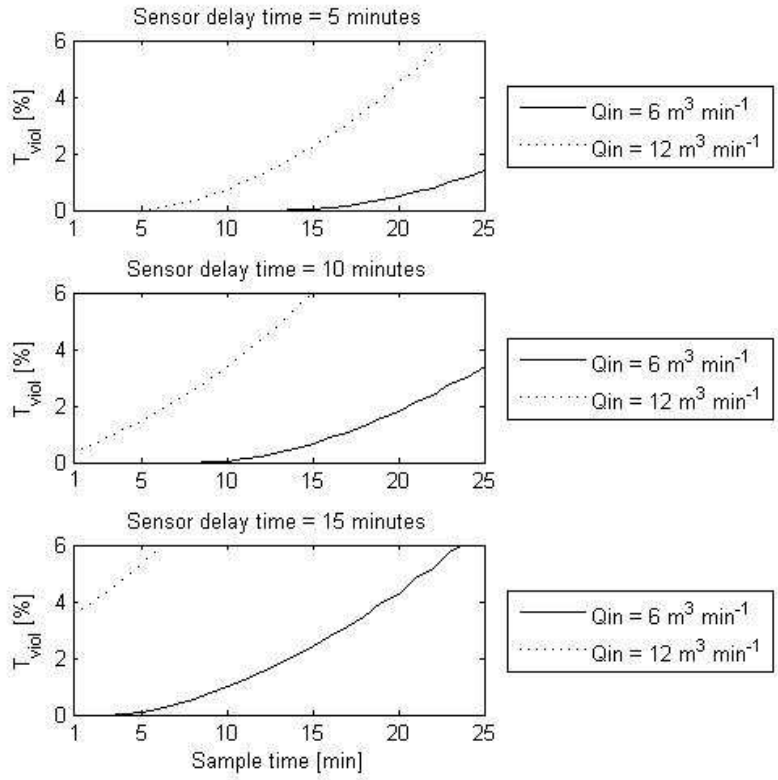

Figure 23: Comparison between the resulting threshold violations when $Q_{in} = 6$ m min^{-1} and $Q_{in} = 12$ m min^{-1}.

The threshold violation limits ($t_{s,lim}$) are the sample times of the points coinciding with the x-axis in Figure 23. If $t_{s,lim}(i)$ is the threshold violation limit for default inflow of a sensor with delay time i minutes, then $t_{s,lim}(5) = 13$ min, $t_{s,lim}(10) = 8$ min and $t_{s,lim}(15) = 3$ min. Correspondingly, $E_{tot,lim}(5) = 70.3$ mJ, $E_{tot,lim}(10) = 114.3$ mJ and $E_{tot,lim}(15) = 304.7$ mJ.

Moreover, the threshold violation level for a sensor delay time of five minutes with doubled inflow is five minutes, with a corresponding E_{tot} of 182.5 mJ. There are no $t_{s,lim}$ values for delay times 10 and 15 minutes, which is seen in the graph.

If the active sample period would decrease, the end of the plume could be detected sooner, which means that the system would dilute the intoxicated sludge basin with incoming nontoxic wastewater sooner. This was tested, by performing an additional performance test with a sensor time delay of 15 minutes, $Q_{in} = 12$ m min^{-1} and by choosing $Asp = 1$ minute. The results from these tests are shown in Figure 24.

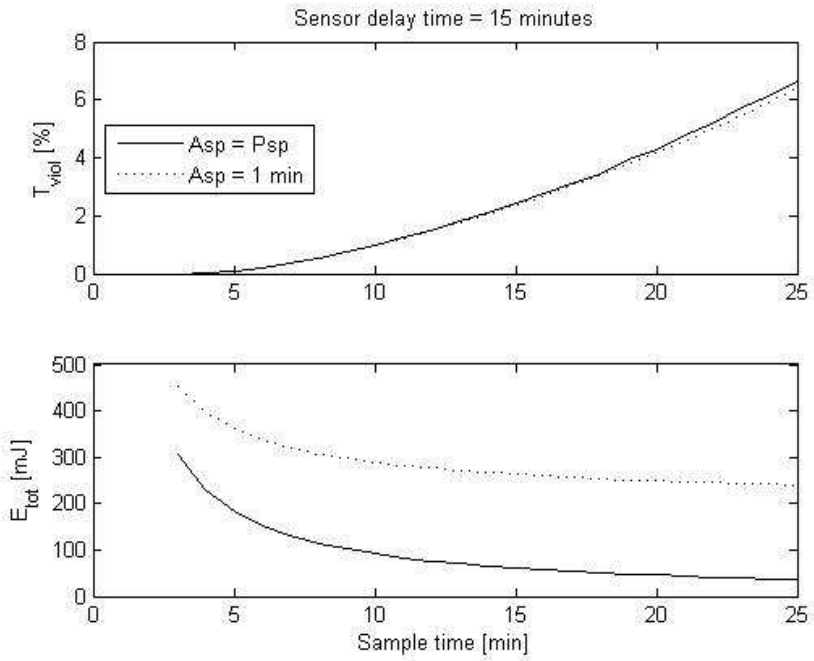

Figure 24: Performance test with $Asp = 1$ minute.

The graphs in Figure 24 show the relative benefits/drawbacks for varying the active sample period. The performance improves slightly by choosing a low active sample period (high active sampling frequency). The energy consumption more than doubles for high sample periods.

Note that Asp does not affect $t_{s,lim}$ since there are no threshold violations to be diluted by nontoxic incoming wastewater when the sample time $= t_{s,lim}$.

6.2 Scenario 2: Two nodes in pipe 3

In this scenario, two nodes are conducted; one at the plant inlet and one 5000 m upstream (in the middle of pipe 3). First, the model was executed with passive and active sample time = 1 minute for both nodes, and the resulting c curves are shown in 25.

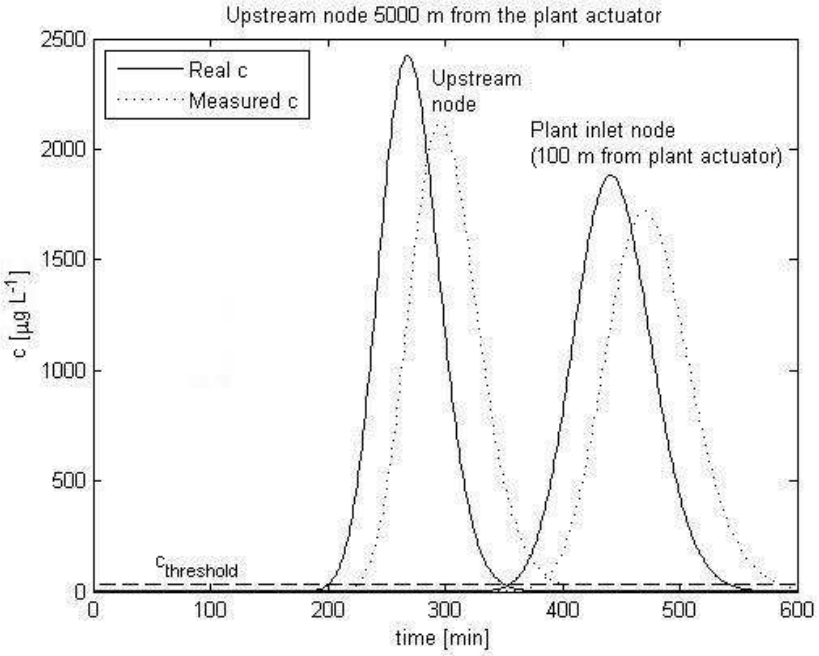

Figure 25: Real c and measured c curves for both nodes.

The time delay $t_{est,D}$ is the time difference from when the upstream node first detects the plume to when the inlet node first detects it in Figure 25. For this particular case; $t_{est,D} = 153$ minutes.

Performance tests were conducted where both nodes always have the same passive sample period, and where the active sample time of the inlet node is 3 minutes. Recall that $t_{s,lim}(15) = 3$ minutes 23, meaning that T_{viol} is guaranteed to be 0 if the alarm signal reaches the inlet node before the toxic plume does. The results are given in Figure 26.

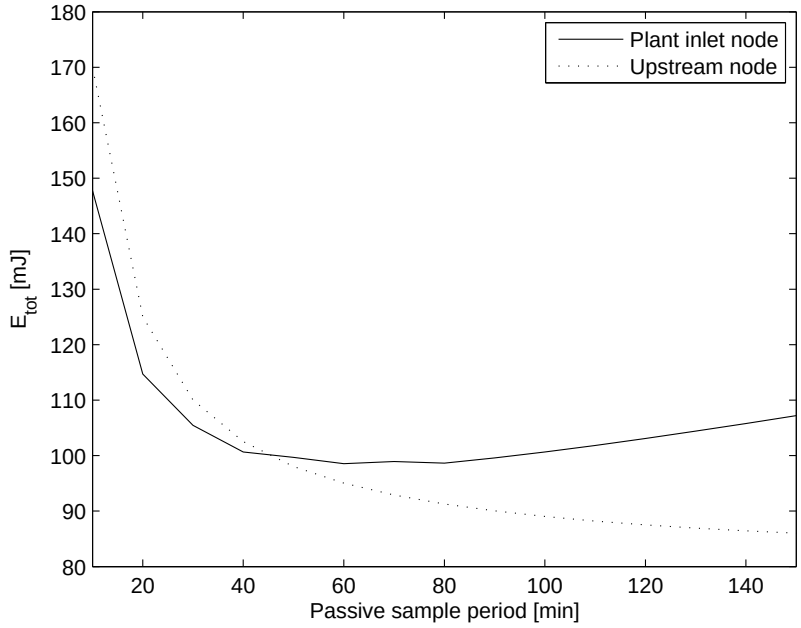

Figure 26: E_{tot} curves for plant inlet node and upstream node respectively.

Recall that $E_{tot,lim}(15) = 304.7$, way above both nodes in Figure 26. Here, the nodes are synchronized (meaning that the time offsets are the same). The node with the highest energy consumption can be seen as the limiting node for the system lifetime, so it is desirable that the nodes energy consumption should be approximately equal.

First, the energy consumption decreases for both nodes with increasing Psp. The upstream node draws more energy than the inlet node because it sends a 32-bit alarm signal (timestamp) 5000 m, which draws a lot of energy. When $Psp \geq 80$, the energy consumption for the inlet node starts to rise. This is because it goes into active mode much earlier in the simulation to compensate for the uncertainty from big sample times.

From Figure 26, one can conclude that it is suitable to chose $Psp = 60$ for this type of sensor, when toxic pulses are expected to come approximately every 900 minutes. If the pulses are expected to come less frequent, it is beneficial to choose $Psp > 60$, because the low sampling frequency will compensate for longer periods in active mode some more. However, if Psp is chosen too big, the plume can do a considerate amount of damage to the activated sludge microbes before it is detected. There is risk of missing the plume if Psp is greater than the time delay of thumb, which is 154 minutes for this case.

To see how the system responds to bad time synchronization, two synchronization tests were conducted according to section 5.4.2. In the first synchronization test, the upstream node is as likely to be ahead of the inlet node as to be behind it, in terms of time offsets. The time offsets differ only 1 minute at max for this synchronization test. The results are shown in Figure 27. In the other test, the upstream node acts as a master node. The results from that synchronization test are shown in Figure 28.

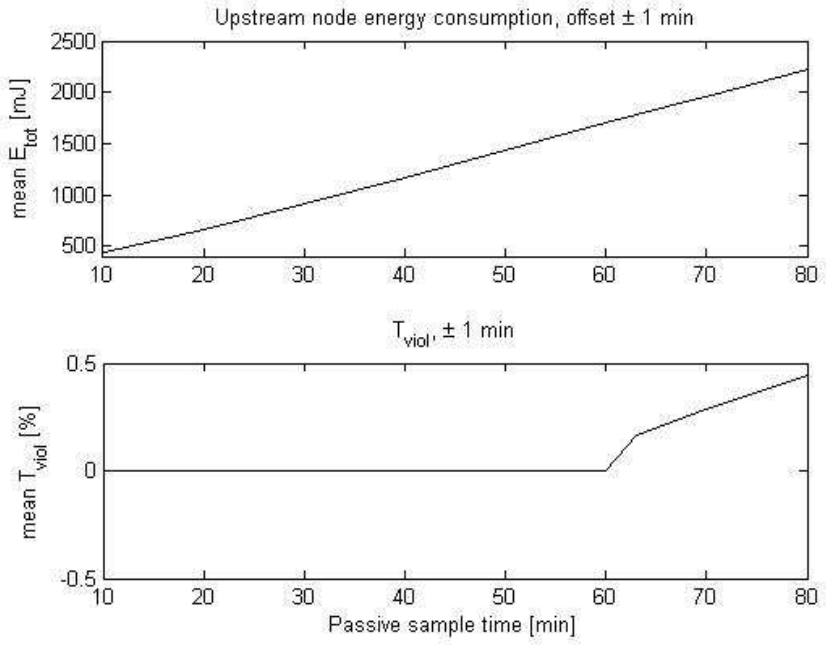

Figure 27: System response with unsynchronized nodes.

The upper graph in Figure 27 shows that this kind of unsynchronization is catastrophic for the upstream node from an energy conservation point of view. The catastrophic results occur because there is a 1/3 risk that the inlet node will be 1 minute ahead of the upstream node, which is equivalent to the case that the inlet node is $Psp - 1$ minutes behind the upstream node. In this case, the upstream node tries to send the energy expensive alarm signal for $Psp - 1$ minutes.

For high Psp, there will even be threshold violations, as shown in the lower graph, since the alarm signal do not reach in time. However, a more realistic scenario is to have the upstream node act as a master node, eliminating the risk of the inlet node being ahead of the upstream node.

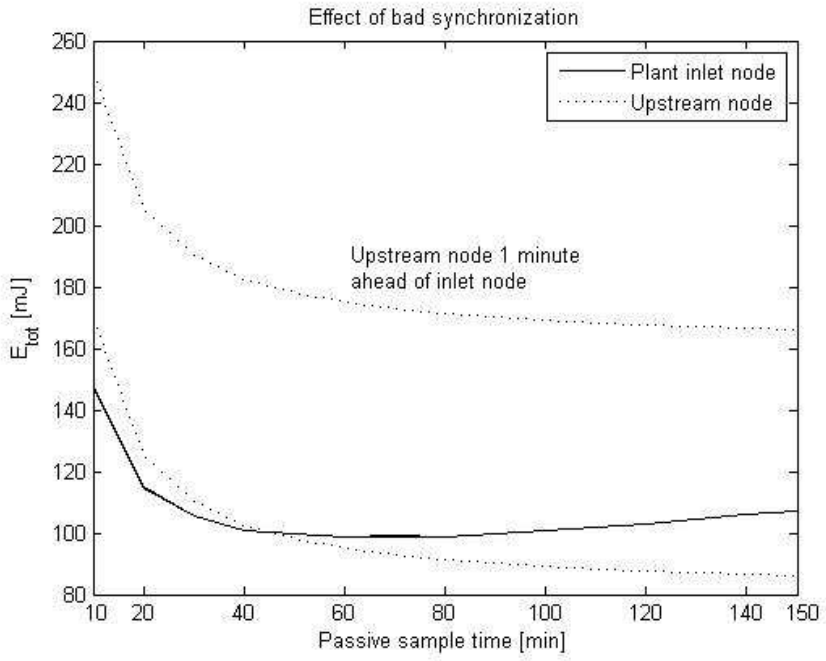

Figure 28: Here are the results of bad synchronization with the upstream node acting as a master node. The two upper curves illustrate the rise in energy consumption from having a 1 minute difference in time offsets.

The energy consumption of the inlet node in Figure 28 is independent of synchronization for time offsets and sample times of these magnitudes. The curve of the upstream node energy consumption is translocated with a magnitude corresponding to the energy cost of the additional alarm signal transfer.

Recall that the energy consumption of the upstream node has to go above 304.7 mJ for the two-node configuration to be pointless, according to Figure 23.

Consider also that these synchronization tests were conducted in a model with a time resolution of 1 minute, which means that messages are transmitted once per minute at max. In reality, the nodes might try to transmit more than once per minute. The real challenge is to have the synchronization gap between the nodes to be less than the transmission time, to avoid wasted transmissions.

6.3 Scenario 3: Two nodes in pipe 3 with variable flow

After 20 minutes of simulation in this scenario, Q_1 and Q_2 doubles, and stays at double default (6 m^3 min^{-1}) throughout the whole simulation time (900 minutes).

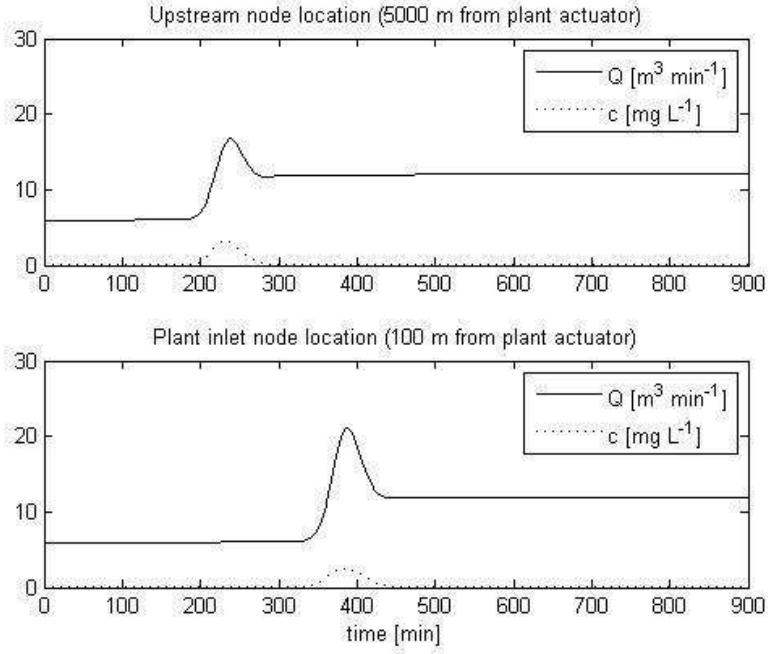

Figure 29: Flow pulse and toxic pulse at the plant inlet.

In Figure 29, the appearance of the toxic pulse and the flow vs time are seen. Note that the toxic concentration is given as mg L^{-1} instead of μg L^{-1} to get an appropriate scaling when comparing to the flow.

During the propagation of the flow step, the flow builds up at the front, then stabilizing at 12 m^3 min^{-1}. This is due to collision of fast and slow water masses (since higher flow indicates higher velocity according to the Manning equation (11)). The propagation of the flow step is described by (19).

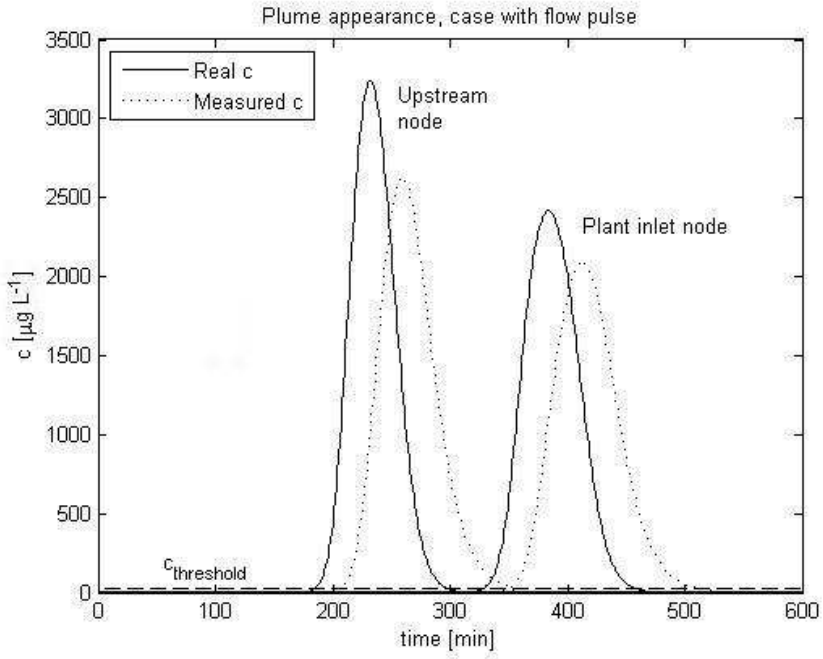

Figure 30: Real c and measured c curves for both nodes.

Figure 30 shows the appearance of the plume at the node sites (which is the same as in Scenario 2) for the case with an inflow step according to Figure 29. Here, both nodes have passive and active sample period = 1 minute. The upstream node is still located in the middle of pipe 3.

For this particular case; $t_{est,D} = 144$ minutes, and $t_{est,in} = 122$ minutes. However, since there are no flow measurements, the plant inflow is expected to be at default (6 m^3 min^{-1}). Therefore, the $t_{est,D}$ used in simulation is the same as in Figure 25, which is 154 minutes (which gives $t_r = 154 - Psp$).

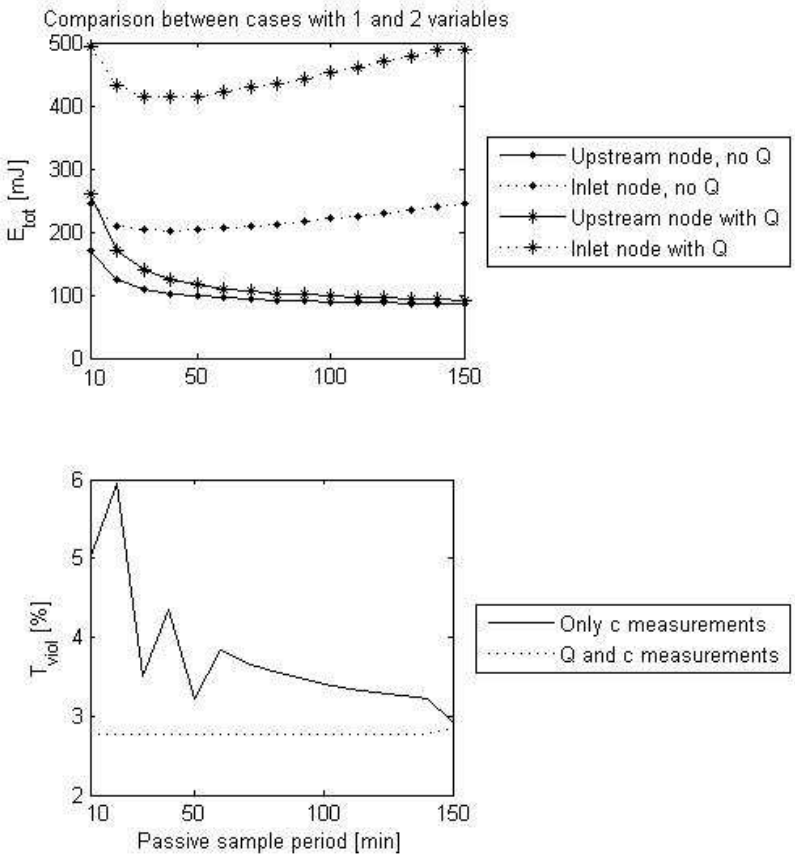

Figure 31: Comparison between cases with 1 and 2 variable measurements. This means that either c is measured only, or both c and Q is measured.

The upper graph in Figure 31 shows the energy consumption of each node in both cases. The more variables measured, the higher the energy consumption. The end node takes the heaviest burden since it is the only node switching to active mode (with $Asp = 1$ minute for this scenario).

Recall that there is no $t_{s,lim}$-value for a node with time delay = 15 minutes when Q_{in} is double default according to Figure 23. This explains why T_{viol} will not become zero, even with a more accurate response time estimation provided by the flow measurements (when flow is measured, $t_{est,D} = 145$ minutes, see Figure 30).

In a case like this, it would be feasible if the upstream node could send its measurements to the actuator directly, so that it could implement the data from the upstream node into its storage strategy. The results of such a node compensation simulation are shown in Figure 32.

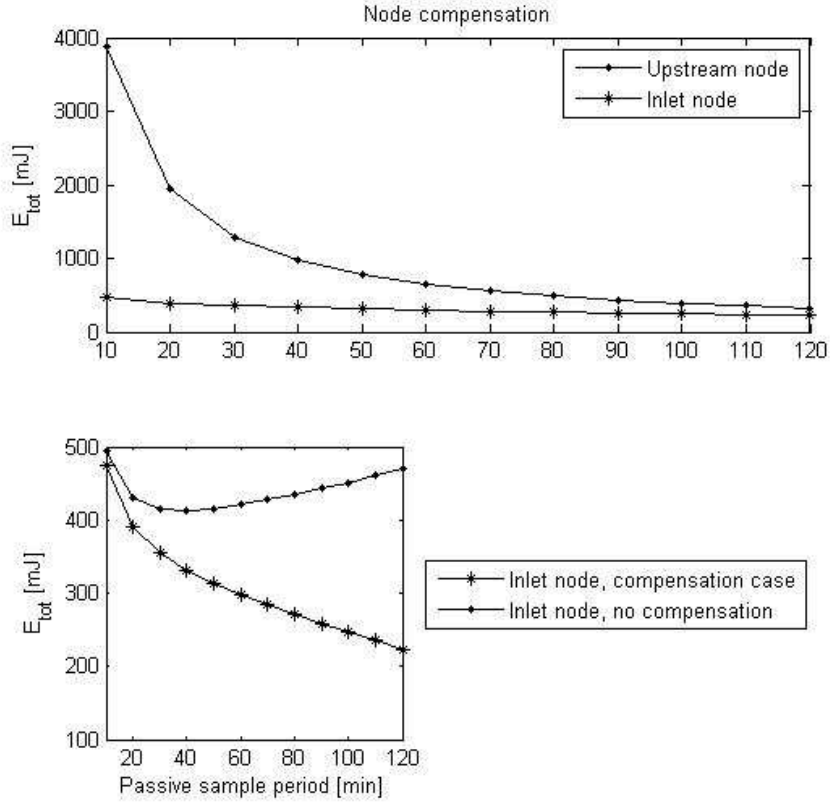

Figure 32: This is the energy consumption graph from the node compensation strategy for sample times $10 - 120$ minutes.

Since both Q and c are measured, $t_{est,in} = 123$ minutes according to Figure 30. Define the *switch response time delay*, or $t_{r,sw}$ as the time delay from when the plume first is detected to when the storage strategy activates due to upstream node compensation. Then, $t_{r,sw} = t_{est,in} - Psp$ minutes.

In the upper graph of Figure 32, the energy consumption of the upstream node get very high, specifically for short sample times. This is because the upstream node send 96-bit data packets (two variables and a timestamp) 5000 m. For this case, $T_{viol} = 0$ % for all sample times. Theoretically, there are risks of threshold violations if $Psp > t_{est,in}$, since the plume could be missed entirely in those cases.

The lower graph shows the difference between energy consumption in the inlet node with or without upstream node compensation. Since the inlet node will not go into active mode until it detects the plume by itself when node compensation is applied, its energy consumption will be lower.

Consider also that the storage tank will likely take in much more wastewater

if the sample period is large. This is because $t_{r,sw}$ is $123 - Psp$ minutes in practice. This means that the storage strategy will execute early after the plume is detected by the upstream node.

7 Conclusions

In this work, a possible application of a WSN was suggested and evaluated in a simulation study. These simulations provided recommendations on sample periods for WSNs with different slowness. These recommendations are referred to as $t_{s,lim}$ in this work. $t_{s,lim} = 13$ minutes if the sensor delay is 5 minutes, 8 minutes if the sensor delay is 10 minutes and 3 minutes if the sensor delay is 15 minutes. This applies to plant inlet nodes with a default inflow of 3 m^3 min^{-1} when the activated sludge basin has a hydraulic retention time of 2 hours. Other dependent factors were the size of incoming toxic pulse and threshold value.

For future work, the sensor delays presumed here should be interpreted in a more realistic manner. They should not only be the product of slow measurements, they could also be a product of simple (but energy cheap) data processing algorithms.

Another conclusion to be drawn is that a configuration of 2 nodes can considerably increase system lifetime because of reductions in energy consumption, compared to a one-node configuration at the plant inlet. However, these advantages can theoretically be drawn from a one-node configuration upstream also, to send alarm signals to the actuator. It is obvious that this configuration lacks redundancy. If there were two nodes, one could take the other node's place if it malfunctions. Another drawback of this configuration is that the storage tank is easily filled up for long sample periods if the storage strategy is on the safe side. These things were not looked into in this work, the only performance limiting factor here was the sensors.

This work also suggests that even a somewhat unsynchronized configuration of one inlet node and one upstream node is better than a one node configuration at the inlet, from an energy perspective. A time offset difference of 1 minute is perhaps more than "somewhat unsynchronized" in reality, but it only results in one wasted alarm signal packet in this model since they are sent once a minute until the notification packet is received.

The conclusions made in this work are preliminary. Before making any definitive recommendations, this scenario needs to be tested in real experiments. Factors not considered here include chemical reactions where toxins are included and different kinds of toxins, for example. What represented one node in this model may in reality represent a cluster of nodes measuring different toxins.

References

[1] C. T. Crowe, D. F. Elger, and J. A. Robertson. *Engineering Fluid Mechanics*, chapter 10, page 409. Donnelley/Willard, 8 edition, 2005.

[2] B. Drafts. Acoustic wave technology sensors. *IEEE Transactions on microwave theory and techniques*, 49(4):795–802, April 2001.

[3] D. Estrin, L. Girod, G. Pottie, and M. Srivastava. Instrumenting the world with wireless sensor networks. In *Proceedings of the IEEE International Conference*, volume 4, pages 2033–2036.

[4] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd Hawaii International Conference on System Sciences*, pages 1–10. MIT, Cambridge, MA, USA, 2000.

[5] Y. Iino and M. Fujita. Wireless sensor network based control system - trade off between sensor power saving and control performance. In *SICE Annual Conference*, pages 2582–2585, 2007.

[6] Y.-W. Lee, S.-K. Ong, and C. Sato. Effects of heavy metals on nitrifying bacteria. *Water Science and Technology*, 36(12):69–74, 1997.

[7] D. Moore, J. Leonard, D. Rus, and S. Teller. Robust distributed network localization with noisy range measurements. In *Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems (SenSys'04)*, pages 50–61. Baltimore, Maryland, USA, november 2004.

[8] C. Oviedo, J. Marquez, and Q. Alonso. Toxic effects of metals on microbial activity in the activated sludge process. *Chemical and biochemical engineering quarterly*, 16(3):139–144, 2002.

[9] T. P. Ruggaber, J. W. Talley, and L. A. Montestruque. Using embedded sensor networks to monitor, control, and reduce CSO events: a pilot study. *Environmental Engineering Science*, 24(2):172–182, 2007.

[10] I. Stoianov, L. Nachman, and S. Madden. PIPENET: a wireless sensor network for pipeline monitoring. In *Proceedings of the 6th international conference on Information processing in sensor networks (IPSN'07)*, pages 264–273. Cambridge, Massachusetts, USA, April 2007.

[11] J. Sundararaman, U. Buy, and A. D. Kshemkalyani. Clock synchronization for wireless sensor networks: A survey. *Ad-Hoc Networks*, 3, 2005.

[12] Y. Tachwali, H. Refai, and J. E. Fagan. Minimizing HVAC energy consumption using a wireless sensor network. In *The 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON)*, pages 439–444, November 2007.

[13] A. N. Tafuri and A. Selvakumar. Wastewater collection system infrastructure research needs in the USA. *Urban Water*, 4, 2002.

[14] A. Ternström. Mikrobiologi och avloppsvattenrening. *Kemisk tidskrift*, 5:18–20, 1992.

[15] T. van Dam and K. Langendoen. An adaptive energy-efficient MAC protocol for wireless sensor networks. In *Proceedings of the First International Conference on Embedded Networked Sensor Systems (Sensys'03)*, pages 171–180. Los Angeles, California, USA, 2003.

[16] G. Wener-Allen, G. Tewari, A. Patel, M. Welsh, and R. Nagpal. Firefly-inspired sensor network synchronicity with realistic radio effects. In *Proceedings of the Third International Conference on Embedded Networked Sensor Systems (Sensys'05)*, pages 142–153. San Diego, California, USA, 2005.

[17] J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. *Computer Networks*, 52(12):2292–2330, 2008.

Personal communication

[18] B. Carlsson. October 2008.