Master Thesis
TVVR 20/5008

Using X-band Radar with a Neural Network to Forecast Combined Sewer Flow

- A case study in Lund

Filip Faust
Per Nelsson

Division of Water Resources Engineering
Department of Building and Environmental Technology
Lund University

Using X-band Radar with a Neural Network to Forecast Combined Sewer Flow

- A case study in Lund

By:
Filip Faust
Per Nelsson

Master Thesis

Division of Water Resources Engineering
Department of Building and Environmental Technology
Lund University
Box 118
221 00 Lund, Sweden

Water Resources Engineering
TVVR-20/5008
ISSN 1101–9824

Lund 2020
www.tvrl.lth.se

Master Thesis
Division of Water Resources Engineering
Department of Building & Environmental Technology
Lund University

Swedish title:	Tillämpning av X-bandradar med ett neuralt nätverk för att prognostisera kombinerat avloppsvattenflöde – En fallstudie i Lund
English title:	Using X-band Radar with a Neural Network to Forecast Combined Sewer Flow – A case study in Lund
Authors:	Filip Faust Per Nelsson
Supervisors:	Magnus Persson, Michael Butts
Examiner:	Rolf Larsson
Language:	English
Year:	2020
Keywords:	X-band radar; neural networks; combined sewer; forecasting; urban hydrology; Lund

Acknowledgements

With this thesis, our five-year journey through the Civil Engineering programme at LTH is coming to an end. Many people have helped us during this spring, and we would like to give our special thanks to:

our supervisors Magnus Persson and Michael Butts, for their patience and input throughout the process,

Nicholas South at VA SYD, for always being engaged and interested in the project. Good luck on your future endeavours!

Lasse Børresen, Charlotte Plum and Peter Rasch at Informetics for answering our questions and providing guidance in our times of need,

Ingrid Nelsson for lending a trained and critical eye to our work,

our comrades in room 1487 that everyday showed truly that the grind never stops,

and Povl Dissing and Benny Andersen for letting us know that soon, the coffee is ready.

Lund, May 2020

Abstract

This study aimed to forecast combined sewer flow into a wastewater treatment plant in Lund, Sweden by using uncalibrated X-band radar data with a neural network. Neural networks have proved themselves useful in the field of forecasting as they can solve multiple kinds of problems and recognise patterns in the data (Alemu *et al.* 2018) as well as model complex real-world problems (Zhang 2012). In 2018, an X-band radar unit was installed in the proximity of Lund which provides precipitation data with high spatial resolution, thus making it suitable for studying precipitation events on a smaller scale (Lengfeld *et al.* 2014). The study concluded that it is possible to accurately forecast combined sewer flow up to 60 minutes ahead of time by only using input variables connected to the catchment of the treatment plant. It was indicated that the prediction time could potentially be extended by adding forecasts of the precipitation as input to the network. The most important input variables were information about the sewage system, a nearby watercourse, the flow at the plant itself as well as information from a rain gauge. The radar is affected by attenuation, degrading the performance of the neural network during large flows.

Contents

1 Introduction

"The goal of forecasting is not to predict the future but to tell you what you need to know to take meaningful action in the present" - Paul Saffo

Since the dawn of time, man has tried to predict and tame the forces of nature and it has proven to be a difficult task. As mankind is dependent on rain, immeasurable efforts have been made throughout history to control this lifegiving asset. Water is not only the source of life, but also a powerful and destructive medium in unwanted places, which is especially relevant in the crowded and urbanised but ultimately fragile cities of today. As society faces increasing threats from volatile weather, modern technology brings great possibilities to open previously closed doors when it comes to predicting the effects of rainfall.

Källby wastewater treatment plant is located in Scania in southern Sweden, and it takes care of the wastewater from the city of Lund and surrounding villages. As parts of the sewer system in Lund consist of combined sewers, the inflow to the treatment plant does not only consist of wastewater but also to large parts of stormwater. This means that in case of a large rain event, the inflow to the treatment plant may increase greatly which is troublesome as the plant operates best at a stable inflow of water to optimise the treatment. For example, a large and sudden inflow of stormwater dilutes the wastewater, reducing the effectiveness of the treatment and disrupts the biological processes (Ellis 2001).

In connection to large rain events occurring several times every year, the flow at the treatment plant reaches very high levels. To avoid problems and ease the operations at the plant, a forecast of the inflow is of great value to know the severity of what is to come. An indication of when the flow is about to increase due to a rain event could also be used to monitor the flow status in the entire sewer system and further warn when there is risk of flooding or combined sewer overflow. The time given by the warning would be key to determine what measures could be taken.

During a precipitation event, the single most important variable deciding the inflow to Källby WWTP is arguably the precipitation. By measuring the precipitation, one could potentially try to forecast large increases in inflow. Traditionally, rain gauges have played an important role in measuring precipitation. As early as 400 years BCE, the people of Arthasastra in India were measuring precipitation with a predecessor to the rain gauge, to determine which kinds of seed they would sow (Strangeways 2010). Although the rain gauge is ingeniously simple, there are certain drawbacks if the purpose

is to use the data for cloudburst analyses. The rain gauge can only measure precipitation at one point but are often used to represent precipitation for a much larger area. The intensity of a strong rainfall can vary greatly in just a few hundred meters, meaning that the value measured in the rain gauge can be misleading for other parts of the area.

An alternative to the rain gauge is a weather radar which measures the precipitation spatially in the air, providing high resolution information of the precipitation as it falls in the area. There are several types of weather radar, for example the long-range and low-resolution C-band radar, which is commonly used in weather forecasting today (SMHI 2017). Another kind of radar is the X-band weather radar, which compared to the C-band radar provides shorter range but higher resolution. This makes the X-band radar data well suited for small scale rain event analysis, for example over a city.

An X-band radar unit was installed in Dalby, 10 km southeast of Lund, in the summer of 2018. After a trial period in the late summer of 2018 it has been collecting data since the spring of 2019 (South *et al.* 2019). The X-band radar produces large quantities of data that requires treatment such as a bias-correction, which makes it inconvenient and time consuming to use with a conventional physics-based model. However, there are other types of models called neural networks, that excel at using large data sets and do not need extensive pre-processing of the data.

Neural networks have become popular in many different kinds of research fields including forecasting, as they provide benefits compared to other conventional methods. Zhang (1998) and Hill *et al* (1993) argue that neural networks could perform well, or even outperform, classical models in tasks related to forecasting. However, the neural network approach is by many still seen as something more experimental in engineering practice and is unfortunately therefore not very common in practical applications. However, this is subject to change as large progress is being made in the field of machine learning, refining the models and allowing use in new applications.

In the context of urban hydraulics, a large advantage with a neural network is that there is no need for a water balance in the system, as it is physics-free modelling. Another benefit of the neural network is that it does not require a bias-correction of the data, as a kind of correction is performed in itself by how the network functions. Additionally, data sources of almost any type can be added to the neural network and the network itself will determine how helpful the input is in improving the results. In conventional models, additional data and parameters need a physical context and expression in the model.

Previous studies using data from the X-band radar in Dalby have been done by Hedell & Kalm (2019), who combined neural networks and radar data for grid points, as well as by Olsson (2019), who used X-band radar data to simulate runoff with a MIKE URBAN model. This project aims to use spatially aggregated X-band radar data for Lund's urban catchment and not for single grid points, as well as using a longer time series of radar data in order to forecast runoff at Källby WWTP. It is desired to evaluate the potential of the X-band radar as it provides great benefits and is a relatively new source of data, which will further be compared with a model using data from rain gauges. The potential to forecast stormwater flow by using only radar data and no additional variables is also of interest as it would make the method applicable to many communities within the range of the X-band radar.

Aim of Project

The purpose of this project is to investigate the potential to forecast the inflow to Källby wastewater treatment plant with a neural network, by using X-band radar data together with other input variables. The project aims to investigate:

- Is the X-band radar data alone enough to make a good forecast?
- Which input variables, network structure and parameters are most useful to improve the quality of the forecast?
- How far into the future can a reliable inflow forecast be made?
- How the forecast compares to a simulated flow made with a conventional model used at VA Syd?

Limitations

To limit the scope of the study, only variables related to Lund catchment are used to train the neural network. Further, the radar data used is limited to the hydrological boundaries of the urban catchment of Lund city. As this study also aims to evaluate the feasibility of training a network with uncalibrated radar data, no treatment on the radar data is performed before it is used to train the model.

2 Theory and Background

In this chapter, the theoretical background to the project is presented. Firstly, the sewage system in Lund and Källby treatment plant are described after which the X-band radar is presented. Lastly, the neural network is introduced.

2.1 Wastewater Management

2.1.1 Wastewater in Sweden

Wastewater and stormwater sewers are important parts of the urban infrastructure, as society is dependent on these functions to prevent waste from accumulating and floods from damaging our cities. In 2015, the public waste- and stormwater pipes in Sweden had a total length of approximately 12 500 km (Lidström 2020). Today, management of both wastewater and stormwater in cities may seem equally important and obvious, but historically stormwater management was the first issue to be solved with planned sewer systems. The increased urbanisation during the 18th century decreased the infiltration capacity for rainfall in larger cities as large areas were made impermeable by structures and roads. During this time there were generally no pipe systems for removal of human waste, which often accumulated in latrines or in the streets (Lidström 2020).

It was not until the late 19th century, as drinking water plumbing was installed in buildings, that wastewater pipes were constructed to remove dirty water containing solid waste from the users. Most often it was directed straight into nearest body of water. During the first half of the 20th century, all new stormwater pipes were directly connected to the wastewater pipes, resulting in the combined sewer systems that still today are common in Swedish cities (Lidström 2020). It was later understood that untreated wastewater had a negative impact on the receiving waters and wastewater treatment plants were therefore constructed from the 1940's and onward. Treatment was initially simple but during the second half of the century, the treatment process was improved with more steps, see figure 1 (Lidström 2020).

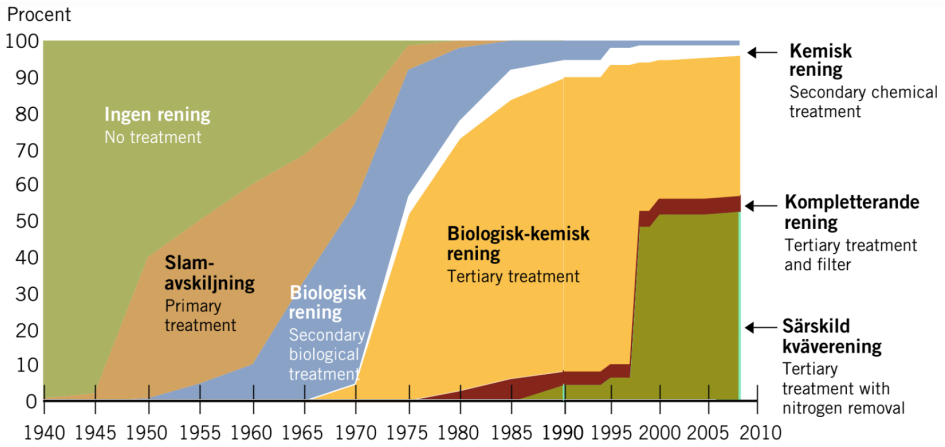

Figure 1: History of wastewater treatment processes in Swedish wastewater treatment plants (Naturvårdsverket 2010).

2.1.2 Combined and Duplicate Systems

The combined sewer system proved to be, and still is, problematic in more than one way. During large precipitation events, where the system runs full, it is not uncommon for combined waste- and stormwater to flow back through the system into basements and streets. The wastewater that did flow to the treatment plants was diluted and in such large quantities that the treatment plant could not treat it properly. To prevent this, combined sewer systems were designed to make combined system overflows (CSO). This allows the untreated water to be discharged directly into the recipient from certain so-called "CSO points" in the pipe system if the water in the sewer system exceeds a certain limit. However, not only large flows are problematic in the combined system but also low or non-existent flows, where the flow is too small to transport solid waste (Lidström 2020).

The solution to these problems was to separate the wastewater and stormwater pipes, i.e. to lead only the wastewater to the wastewater treatment plants in pipes with smaller dimensions to ensure a relatively constant flow while the stormwater is led in a separate system that transports it directly to the recipient. This type of system is called a duplicate system. Another solution was to construct retention basins on the combined systems where large volumes of water could be retained and later released to evenly spread out the load on the wastewater treatment plants (Lidström 2020).

From the 1950's and onwards, the duplicate system has been the dominating system built in newly developed areas. In areas where stormwater is not a large issue, such as in smaller settlements and rural areas, separated sys-

tems are also common, which is essentially a duplicate system but without any underground stormwater pipe system (Lidström 2020). Today, existing combined systems are systematically being replaced with duplicate systems in most cities, but there are still many combined systems in Swedish cities.

2.1.3 Inflow and Infiltration

Apart from the water that the pipes are designed to contain, such as wastewater in separate wastewater systems and wastewater plus stormwater in combined systems, there is always an addition of unwanted water. This water is referred to as infiltration/inflow or I/I. Infiltration is the process of groundwater entering the sewer pipe, for example thought cracks and joints. Inflow is when rainwater enters the sewer pipe directly, for example through leaks in manhole covers (Karpf and Krebs 2011). In combined systems, inflow can be considered a part of the normal function of the pipe as they are supposed to also handle storm- and drainage water. Figure 2 shows I/I in principle for a separate wastewater (sanitary) sewer.

The severity of I/I is determined mostly from the condition of the pipe, where an old and worn pipe will be more susceptible. In wastewater pipes, I/I can account for at more than 50 % of the normal dry weather wastewater flow (Svenskt Vatten 2016). The total amount of I/I in a sewer system is determined by the condition of the pipes, the total length of the sewer system and the groundwater level. If the groundwater level is close to the ground surface or above the pipe depth, the infiltration will increase.

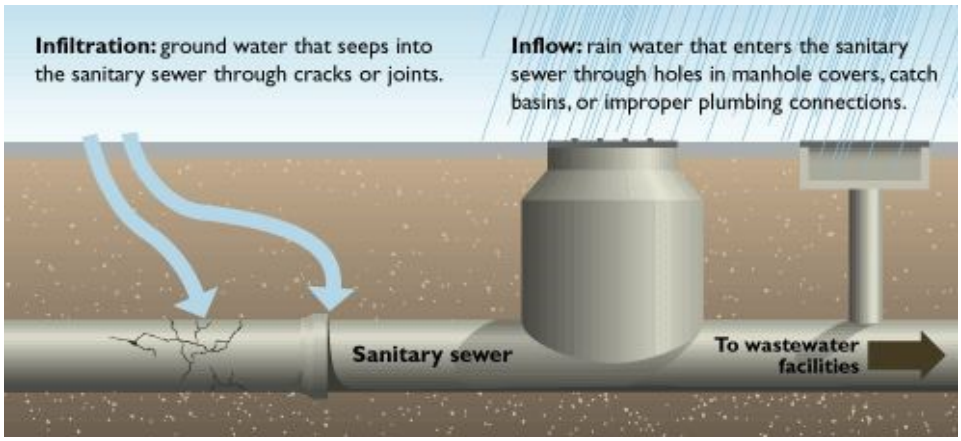

Figure 2: Definitions of infiltration and inflow in a wastewater pipe (City of Salem 2020).

As I/I increases the flow of water in wastewater sewers, the hydraulic loading on wastewater treatment plants is also increased. This decreases the efficiency of treatment as the wastewater gets more diluted from the often relatively clean groundwater or rainwater. The extra flow from I/I also accelerates pipe deterioration and the infiltration can cause erosion around the pipe, destabilising the soil and endangering nearby structures (Karpf and Krebs 2001). However, some effects of I/I can be considered beneficial. For example, old and leaky pipes drain the soil of groundwater, keeping the groundwater level from rising and flooding basements (Gustafsson 2000). The extra flow can also help against sediment build-up and smell development in the pipe system.

2.2 Lund

2.2.1 Wastewater in Lund

Källby wastewater treatment plant (WWTP) is located in the southern parts of Lund, Sweden, and treats the wastewater from large parts of Lund municipality, see figure 3. Apart from Lund and Dalby, which both have areas with combined sewer systems, wastewater from Veberöd, Genarp and Björnstorp reaches Källby WWTP. The treatment plant is operated by VA SYD, an organisation responsible both for distributing drinking water and taking care of wastewater in southern Scania. The organisation is a collaboration between the municipalities it operates in, which are Lund, Malmö, Burlöv, Lomma and Eslöv.

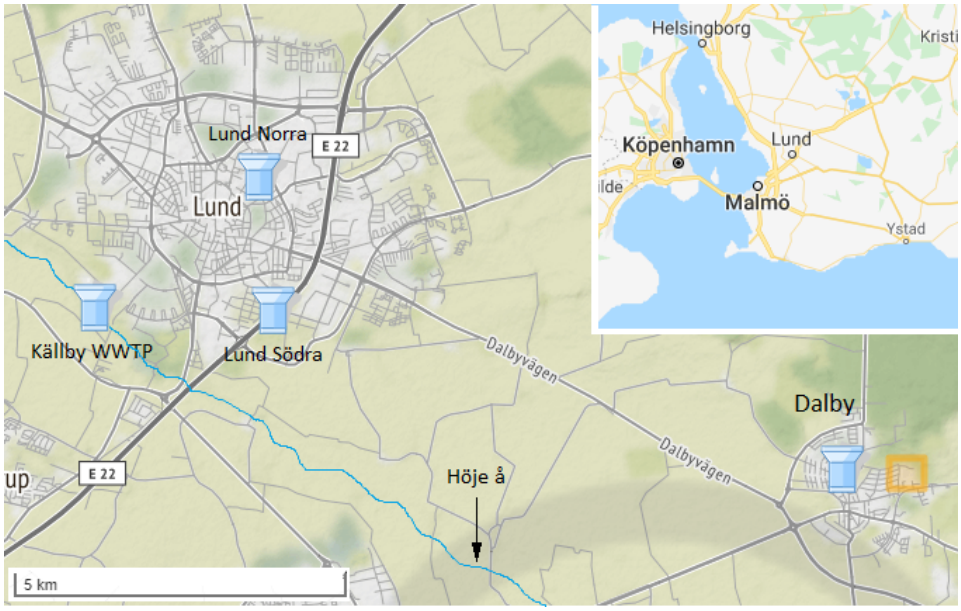

Figure 3: The location of Lund in the southernmost part of Sweden, close to both Malmö and Copenhagen, with Dalby located roughly 10 km southeast of Lund (Google and VA SYD 2020).

The sewer system in Lund consists of both separated and combined pipes, where the combined pipes make up roughly 10 % of the total wastewater system. The combined pipes are only present in the old central parts of Lund, where they make up as much as 84 % of the pipe system (VA Syd 2018c). Due to elevation variations in the city, there is a natural beneficial gradient towards Källby wastewater treatment plant, which is located at a low elevation in the south of Lund. Basement floodings linked to the combined system have not been a large issue in Lund, but in some parts, they can be caused by I/I (VA Syd 2018c). Figure 4 shows the extent of the combined system in Lund.

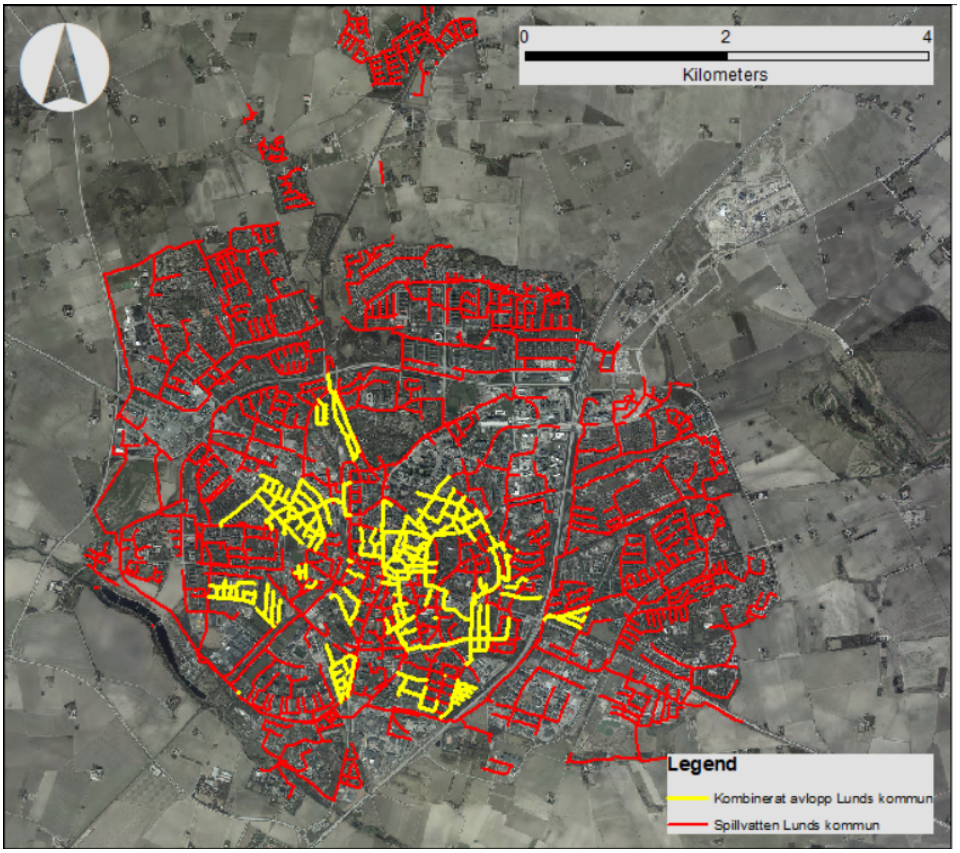

Figure 4: Map of Lund with sewer pipes. Combined pipes are yellow and separated pipes are red.

Since the sewer network in Lund is a mix of duplicate and combined systems, the inflow to the treatment plant consists of both waste- and stormwater. The flow of pure wastewater is normally rather consistent, even though it varies over the day, but the stormwater flow may vary greatly depending on the occurrence of rain events. The main weakness of the sewer system in Lund is the additional stormwater that puts a heavy load on the system risking basement floodings and CSOs, as well as putting heavy hydraulic loads at the WWTP. All CSOs in urban areas in Lund may eventually end up in the recipient Höje å, which has been classified as very sensitive to pollution as well as sensitive to increases in flow and nutrients (VA Syd 2018c). Sudden fluctuations or increases in incoming water to the wastewater treatment plant can further wipe out nitrifying bacteria which cripples the treatment, resulting in poorly treated water in the outflow requiring hard and careful work to restore the biological treatment processes (Bashide 2015).

A measurement campaign of wastewater flows was carried out in Lund by VA SYD in the late summer of 2018. The hydrodynamic flow in wastewater pipes was investigated in 12 points all around Lund. In all points a diurnal flow pattern could be observed during dry weather, as can be expected from wastewater pipes. However, during precipitation events, the flow in all measuring points showed a sharp increase followed by a rapid decline. This was interpreted as a clear indication that runoff from hard surfaces is incorrectly connected to the wastewater system near these points (VA Fälttjänst AB 2018).

VA Syd has defined hydrological boundaries for Lund city, meaning that theoretically only water that falls inside these polygons will end up at Källby WWTP, via the wastewater system through I/I. Lund is further divided in several sub-catchments, depending on how surface runoff is routed to the wastewater sewer system. These boundaries are illustrated in figure 5.

Figure 5: The hydrological boundaries of sub-catchments in Lund regarding runoff to the wastewater sewer system.

2.2.2 Potential Relocation of Källby WWTP

As the city of Lund is expanding, Lund municipality has plans for development in the southwest areas of Lund, where Källby WWTP is located. There are additional plans for a new train station in the vicinity to be constructed within a few years (Lunds Kommun 2018). The future of Källby wastewater treatment plant is currently being discussed, and two options are currently on the table. Within a few years, it must either be rebuilt and upgraded, or demolished, requiring the relocation of the treatment to Sjölunda wastewater treatment plant in Malmö (VA Syd 2019a). Further, the receiving capacity of Höje å, where the treatment plant discharges its water, is not enough to handle an increasing load from Källby (VA Syd 2019b). If the treatment is relocated to Malmö, the wastewater would be transported from Lund to Malmö in either a pipeline or an underground tunnel. In that case, the forecasted inflow to the tunnel would theoretically be the same as the inflow to Källby WWTP, which is investigated in this thesis.

2.3 Weather Radar

2.3.1 Overview of Weather Radars

Whereas a conventional rain gauge measures the accumulated volume of rain in the gauge, a radar measures the reflectivity of rain particles in the air (Einfalt *et al.* 2004). The radar unit rotates and scans the air in 360 degrees at several inclinations, giving an almost full picture of the atmosphere (South *et al.* 2019). There are different kinds of radars used for measuring precipitation, mainly C-, S- and X-band radars, all of which have their certain properties. The C- and S-band radar are commonly used in large scale weather forecasting because of their long range, exceeding 100 km (Einfalt *et al.* 2004).

2.3.2 X-Band Radar - A High Resolution Radar

The X-band radar provides high frequency and high-resolution data in a range of 20 – 100 km with a spatial resolution as low as 20 m (Lengfeld *et al.* 2014). Compared to the C- and S-band radars, the X-band radar gives more detailed information within a shorter range. C- and S-band radars have a significantly longer range at a lower spatial resolution, which is roughly 0.25 – 2 km for C-band and 1 – 4 km for S-band (South *et al.* 2019). The X-band radar serves as a good complement to the C- and S-band radars, especially while studying local precipitation events that can vary greatly in just a few hundred meters where the resolution of the C- or S-band radar is not satisfactory (Lengfeld *et al.* 2014). Another benefit of

the X-band radar is the resolution in time, where the X-band radar provides up to 1 update per min and the C- and S-band radars take roughly 5 – 15 min between the updates (South *et al.* 2019).

Since the radar measures reflectivity from particles in the air, one drawback is that the radar signal can get completely blocked in case of a high intensity rain with a high density of particles. This phenomenon is called *attenuation*, meaning that the radar fails to register potential precipitation behind an area with heavy precipitation (South *et al.* 2019). Using several radar units in a network, where one radar might register precipitation that another one misses, would reduce clutter and improve the result (Lengfeld *et al.* 2014). To improve the performance of the X-band radar it can further be calibrated together with a C- or S-band radar, which operates at different frequencies not as affected by attenuation (Lengfeld 2016).

2.3.3 Installation of Sweden's First X-band Radar in Dalby

In the summer of 2018, an X-band radar unit was placed on the water tower of Dalby, illustrated in figure 6), with the purpose of analysing rain and cloudbursts in the region (VA Syd 2018b). Dalby is located roughly 10 km south east of Lund and on a high ground, making it an ideal location for a radar as it provides great vision, clear of most obstacles, over Lund and the rest of southern Scania. The radar unit is a cooperation between VA SYD, LTH, Sweden Water Research and SMHI and is the first X-band radar to be placed on Swedish soil. South *et al* (2019) argue that there is great potential in investing in more X-band radars in Scania and to potentially cooperate with Danish counterparts to provide full coverage of the region.

Figure 6: Dalby water tower on top of which the X-band radar unit is installed (VA Syd 2018b).

The X-band radar unit in Dalby is a Compact Dual Polarimetric X-band Doppler weather radar WR-2100 and has a range of 60 km with a maximum resolution of 50 m. It can rotate at a maximum of 16 revolutions per minute and has a frequency of 9.4 GHz (Furuno 2020). The radar measures precipitation on 4 different angles above horizontal level, where level 1 corresponds to 2 degrees, level 2 to 4 degrees, level 3 to 8 degrees and level 4 to 10 degrees. The radar unit in Dalby operates with a measuring resolution of 500 m, providing data arranged in a grid with cells of this density (South *et al.* 2019).

2.3.4 Measuring Precipitation with Rain Gauges and Radar

There are three rain gauges placed in Lund that measure precipitation, shown in figure 3. The rain gauges are so called *Tipping Bucket* gauges, which register a value for every 0.2 mm precipitation that they have measured. Because of their simplicity, rain gauges have for a long time played an important role in measuring precipitation. However, there are certain drawbacks with the rain gauges. As they only measure precipitation at a single point, the value measured in the gauge becomes representative for the entire area around it. This can be problematic, as the intensity of a rainfall can vary greatly in just a few hundred meters.

An example is illustrated in figure 7, where an intense rainfall completely missed one out of three rain gauges as it passed over the city of Trelleborg

in the summer of 2018. By only using information from the eastern gauge in the figure, it would seem like there had not been any precipitation falling in Trelleborg. Additionally, regular maintenance of the rain gauges is also required as there might be things such as leaves or dead animals physically blocking the inflow to the rain gauge, which could yield inaccurate readings.

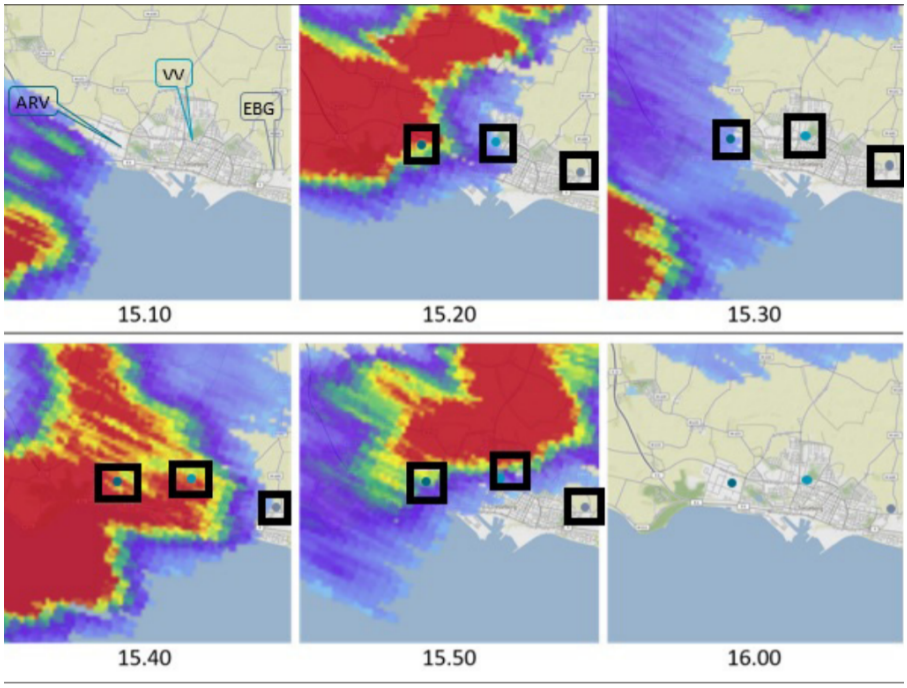

Figure 7: An intense rainfall passing by Trelleborg in southern Sweden 2018-08-11. Out of the three marked rain gauges, the rainfall completely missed the eastern one (South et al., 2019).

A radar offers solutions to some of the problems that a rain gauge cannot handle very well. While a radar is far more complex than a rain gauge and needs extensive calibration, the radar measures the precipitation spatially which is a great benefit compared to the rain gauge. The radar can also measure precipitation at any point within range without needing a rain gauge or similar reading instrument at the location of the rain. However, a rain gauge is a cheap and simple alternative to a radar unit, making it a good option for users who cannot access a radar. Most importantly, the rain gauge is not affected by attenuation effects. Thus, a rain gauge can give valuable complementing information when attenuation effects reduce the quality of the radar estimates. The rain gauge and radar should necessarily not either be seen as competitors, but rather as two different tools that can work well together compensating for each other's weaknesses.

2.4 Neural Networks

2.4.1 Principles of a Neural Network

A neural network is a data driven model that treats information in a process that tries to mimic the human brain, in a sense that nodes are used for building a model, where the nodes are similar to the neurons in the human brain. This approach enables the model to solve multiple kinds of problems and recognise patterns (Alemu *et al.* 2018). The model excels in treating large amounts of data and the network is trained to optimise the ability to produce a result. The optimisation process is complex and there are several different algorithms developed to determine the error, called *loss*, all in order to change the weights connected to the nodes in the model (Zhang 2012). The network is trained on historic data of what it tries to produce, called the *target*.

Generally, a network is developed with a training phase that includes validation and an independent test period after to verify the output of the model. This is slightly different from a conventional model but can be compared to the division into a calibration and a validation period, where the training corresponds to the calibration and the testing to the validation. The validation in the training period is done automatically on a pre-selected portion of the data as part of the training process, to evaluate the output of the network (Zhang 2012). Testing is done on a separate set of data after the model is finished training.

If the input to the network is changing over time, the model can adapt to the changes, which is a great benefit that comes with using a machine learning method. A neural network can vary in depth, ranging from deep to shallow. The depth corresponds to the number of hidden layers in the network, where a shallow network may have only one hidden layer, but a deep network can have many (Hardesty 2017). A deep network allows for a sophisticated structure to approach complex problems.

2.4.2 Architecture of the Network

The neural network consists of different layers which are made up of nodes. There is an input layer, an output layer and hidden layers in between those, illustrated in figure 8. The nodes in the input layer recieves the input data and sends it further to the nodes in the hidden layers, which have weights attached to them. If the network finds one input variable of low importance, the weight attached to this variable will be low. The final hidden layer forwards the data to the output layer, which is the last and final layer. The output layer then tries to predict the most likely output (Zhang 2012 and

Solomatine *et al.* 2008).

The point of the neural network is that the network should learn and develop from previous experience. The learning experience is seen as change in weights in the hidden layer, as the network learns and is optimised (Gurney 1997). This means that the hidden layer is tweaked and iterated until the output is considered most satisfactory. There is generally only one hidden layer in neural networks used for forecasting, which in several studies have proved to be sufficient. However, other researchers find that using two layers is beneficial compared to using one layer for the purpose of forecasting (Zhang *et al.* 1998).

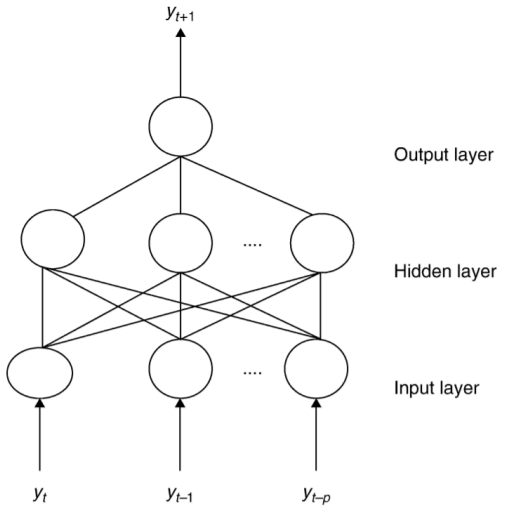

Figure 8: Principal structure of a neural network model (Zhang 2012).

The number of input nodes often corresponds to the number of input variables to the network. Choosing good input variables is crucial if the network should be able to recognise patterns in the data, making it one of the most important factors in the network (Zhang *et al.* 1998). Although the network itself should be capable of determining which input variables are relevant and not, feeding the network irrelevant data may degrade the result.

The number of nodes in the hidden layer is also a parameter to consider while designing the neural network. Networks with a low number of nodes in the hidden layer have a good ability to generalise, but models with too few nodes may not be able to capture and recognise the patterns of the data, which is crucial for the network (Zhang *et al.* 1998).

2.4.3 Parameters in the Neural Network

The loss function is a function displaying the error in the model, measuring the difference between the predicted output against the actual value of the target (Wu *et al.* 2019). The loss function depends on the values of the weights and consists of several local minima and a global minimum. The global minimum corresponds to when the error in the model is at lowest, meaning that the weights for the nodes in the hidden layer are most optimal.

The learning rate is connected to the loss function as it is the parameter deciding the pace finding the minima (Wu *et al.* 2019). Deciding the optimal learning rate is difficult and not always a straight-forward process, as it varies from case to case. If the learning rate is too high, the model may find a local minimum but not the global minimum, tricking the model that the local minimum is the global minimum and the best solution. On the other hand, a low learning rate will slow down the training process (Wu *et al.* 2019). Figure 9 displays how the learning rate can affect the iteration process for finding the most optimal weights.

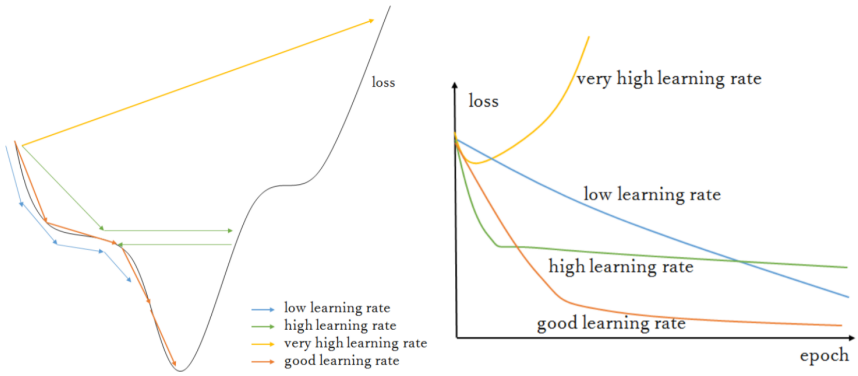

Figure 9: Examples of how the learning rate affects the loss (Zhang & Song 2019).

Activation functions, also called transfer functions, determine the total weighed sum in a node, and produces an output for the node (Nwankpa *et al.* 2018). If the sum of the node is under a certain threshold value, it can be chosen to not be used or sent forward in the network. There are different kinds of activation functions, linear and non-linear, which all have their benefits and drawbacks depending on the area of interest. A common type of transfer function is a sigmoidal function, which has the shape of a tangent function and gives an output value between 0 and 1, depending on the total value of the input to the node. Other transfer functions can be binary, meaning that they yield either a value of 1 or 0.

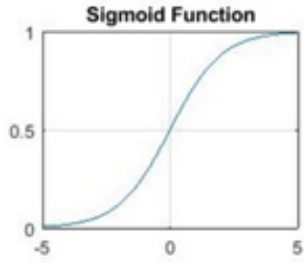

Figure 10: A sigmoid activation function (Nwankpa et al., 2018).

Another common activation function is a rectified linear unit ($ReLU$). The ReLU function yields an output of 0 for all input values below 0 and yields a linear function for all values over or equal to 0, see figure 11 (Agarp 2018).

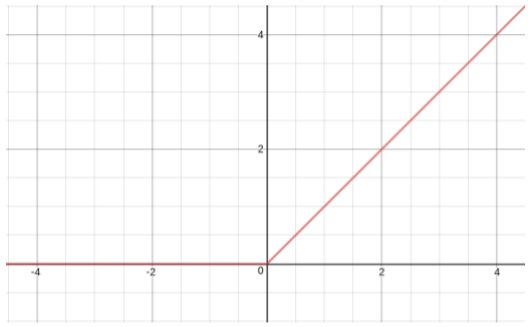

Figure 11: Rectified linear unit (ReLU) activation function which produces a linear output for input values where x >0 and 0 for all other.

To train and test the network, the input data is divided into two phases, usually referred to as *data division*, a training period and a testing period. The division of data is something that will greatly affect the evaluation of the performance of the model. There is no consensus regarding how data division should be done, but a general guide mark is to use roughly 80 % of the data for training and 20 % for testing (Zhang *et al.* 1998). Zhang (2012) suggests that both phases should individually be representative for the data as a whole and that the division should be done randomly in order to make it so.

The training period further includes a validation period which makes up a smaller part of the total training period. The purpose of the validation period is to continuously evaluate the performance of the model during the training and select a final model which is to be used (Zhang 2012). The best performing model is selected by the network to be used for the independent test period.

2.4.4 Problems with Machine Learning

While machine learning certainly has its benefits, it also comes with potential drawbacks such as overfitting and problems with interpretation of the results. Overfitting is a common problem occuring while using a neural network, meaning that the model fails to recognise the patterns in the input data. The model further adapts too much to the training data, making the model perform poorly when it has to predict an output on its own (Chicco 2017). Figure 12 illustrates the problem with overfitting well for a polynomial model, where an order over a certain number starts to decrease the performance of the model.

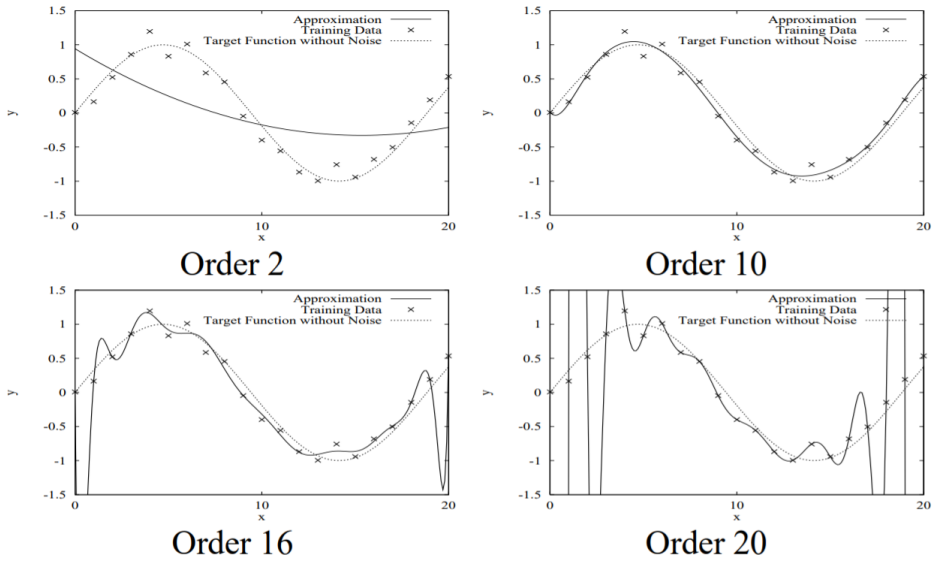

Figure 12: How different orders affects the approximation and how the result degrades as the model is overfitted (Lawrence et al., 1997).

Zhang (1998) proposes that overfitting can be avoided by restricting the number of nodes in the hidden layer and concludes that several other studies have shown that networks with the same number of nodes in the hidden layers as in the input layer performs good in terms of overfitting. Further, Ying (2018) concludes that overfitting cannot be completely avoided, but algorithms that try to stop the training before it gets overtrained can be introduced to the network to minimise overfitting.

2.4.5 Difference Compared to a Physics-based Model

The neural network differs greatly from a physics-based rainfall runoff model. A physics-based model like MIKE URBAN must respect the physical laws and relationships while keeping a water balance (DHI 2017). A neural network model can somewhat neglect this as it tries to make the best prediction possible using the input data it has available by recognising the connections between the input data and target output. As the physics-based model is versatile and has been used for a long time, they are today a common planning and forecasting tool.

The differences between the two kinds of model gives them both their own characteristics, strengths and weaknesses. The neural network benefits from that it does not need to use bias-corrected data, as a kind of bias-correction is performed in how the network functions. Additionally, as the neural network completely neglects the water balance, new sources of data input can easily be added and investigated as they do not need to be implemented into the model software.

However, a water balance can be considered a large advantage to the physics-based model, as it controls the model and makes it more robust and trustworthy. This can be compared to the neural network, which is a black box model, where no transparency in the operations is given. Further, if something significant is altered in the sewer network, such as adding a large detention basin to a part of the system, this could be added as a new component to the system in a MIKE URBAN model. As a neural network model is trained with historical data, a significant change in the real system could potentially make the model obsolete. Training a new model requires data, which could take years to record if there is no data.

3 Method, Model and Data Treatment

In this section, the method of carrying out the project is described. Firstly, the neural network used is described after which the quality of the radar data is evaluated. The different cases used in the project are described as well as the input variables tested. Finally, the investigation of prediction time and comparison with a conventional model is described. Further, the forecast is also referred to as a *prediction*.

3.1 The Model and Output

3.1.1 The Neural Network

The neural network in this project is Python-based in the platform *Tensorflow* and is developed by the company *Informetics*. The network uses a learning rate of 0.1 and a ReLU activation function, which is considered a suitable activation function as it is cheap in terms of computation and handles negative input in an efficient manner. The model has one hidden layer with 8 nodes and makes a forecast 60 min ahead of time if nothing else is specified. The prediction output is the mean of many predictions in form of a gaussian distribution for each time step. The prediction mean is the value used as the official prediction value of the model. The standard deviation for the gaussian distribution is also produced.

Certain features are built into the script to improve the performance of the model. Sine functions with periods corresponding to the length of 1 day and 1 week are automatically implemented as input with the purpose to represent patterns of these durations, thereby assisting the model. These are supposed to simulate the daily and weekly patterns that the wastewater has. For example, the base flow of wastewater tends to follow a diurnal pattern, which could be compared to a sine curve with a 1-day period.

Training several neural networks with the exact same dataset and parameters will not produce identical models. This is because networks learn slightly differently during the training process. The differences between the models are usually small, but it is important to acknowledge that these internal differences between models exist when drawing conclusions from comparisons between different tests and cases. An example of this can be seen in Appendix A - Visualisation of variation between identically trained neural networks.

If there is data missing for any input variable for a certain timestamp, all other inputs for the timestamp is automatically removed, even if there are data for the other ones. This allows the neural network to be trained despite

incomplete time series.

3.1.2 Optimisation of Network Architecture

As both the input variables and the structure of the network affect the forecast, it is desired to pursuit the best possible configuration of the parameters in the network. The network structure is studied to evaluate its importance compared to the input variables. The parameters in the network that are studied are the number of hidden nodes, the number of hidden layers and the *roll durations*. The evaluation of the parameters is done entirely by visually comparing graphs showing the performance when parameters are altered.

Roll durations is a setting in the configuration of the neural network where one or several rolling averages of the input variables can be added. This means that the average value of every input variable for the selected roll duration is added as an extra input variable. The purpose is to implement a memory of past events as inputs, which can represent for example increased levels in base flow in the sewer network or the degree of saturation in the soil. In all tests, unless otherwise stated, the roll durations are set to 15 min, 1 h, 3 h and 8 h. For example, the roll duration for 15 min adds an extra variable with the average value for the last 15 minutes for every other input variable. The rolling averages are variables in the model, but since the setting for how many and how long these should be is set in the configuration of the neural network, they are presented together with the parameters.

3.1.3 Selection of Training and Validation Period

A long time series of data is desired to give the network the best conditions to train and validate. In this case, the radar data is the limiting factor as the radar was not continuously measuring precipitation until the beginning of May 2019. The radar had previously only occasionally been active from its installation in the summer of 2018. This means that the time series that is used in this project stretches from 1 May 2019 until 23 Jan 2020.

The training period is selected to May - Nov 2019 and the validation period includes Dec 2019 - Jan 2020. No testing is performed in this study, as the training and validation is considered enough for the purpose of investigating the influence of different variables and parameters on the model performance.

3.2 Status and Quality of Radar Data

3.2.1 Identified Issues

There are several different known problems with the radar unit in Dalby and its performance. The radar unit has been inactive for several periods of time, thus not measuring any precipitation for these periods. Additionally, the radar suffers from other issues, potentially clutter, resulting in among other things, overestimation of precipitation. Because of this, a clutter filter has been installed, completely removing data for certain areas of Lund. It is desired to know where and when the issues affect the radar in order to improve and adapt the performance of the model.

A previous study by Kalm & Hedell (2019) indicated that the level 2 angle of the radar beam is the level with most useful radar data for runoff modelling purposes. However, it was not possible to retrieve radar data for level 2 for the studied period of time. Instead, the data for level 3 is used. Level 2 and level 3 corresponds to a radar beam 4 respectively 8 degrees over horizontal level.

3.2.2 Inactive Periods

As the radar unit has been inactive for several periods of time, there are gaps in the time series of radar data. The periods of downtime are presented in table 1. For some of the periods there was precipitation recorded by nearby rain gauges. For periods without any registered precipitation by nearby rain gauges, the gaps in the radar time series will be set to 0, as it is assumed that no precipitation was falling during this time period. In the 26-day gap during the summer of 2019, precipitation was only recorded from 21/7 – 31/7 but not for the remaining days. The data for the remaining days in the gap is thus set to 0.

Table 1: Some of the longer downtimes for the radar in Dalby. The rain gauge refers to the gauge at Källby treatment plant.

Start	**Stop**	**Downtime**	**Rain recorded by rain gauge**
2019-06-12 19:53	2019-06-12 23:59	4 h	Yes
2019-06-29 00:00	2019-06-30 23:59	2 days	No
2019-07-13 00:00	2019-08-07 23:59	26 days	Yes
2019-09-15 00:00	2019-09-15 23:59	1 day	No
2019-09-21 00:00	2019-09-24 23:59	4 days	Yes
2019-10-01 00:00	2019-10-02 23:59	2 days	Yes
2019-10-22 00:00	2019-10-22 23:59	1 day	No
2019-11-08 00:00	2019-11-26 23:59	18 days	Yes

Additional periods with problems were discovered when retrieving radar

data for certain areas. The reason is unknown, but the data could not be retrieved, and the result was periods of varying durations lacking radar data. In total, around 15 % of the radar data was lost for the study period due to various reasons.

3.2.3 Spatial Distribution of Errors

The radar measures the quantities of precipitation vastly differently over the city. Therefore, it is desired to investigate how different parts of the city are affected by the errors related to the radar unit. A grid is established with the purpose to determine the quality of the radar data in each grid point. This is done by accumulating and comparing the total measured precipitation for each grid point. The accumulated precipitation measured by the radar can then be compared with the accumulated rainfall from the three rain gauges that are placed around Lund, see figure 3. Ideally, the precipitation measured by the rain gauges and the radar follow the same pattern. As seen in figure 13, the three rain gauges follow the same pattern which indicates that they measure roughly the same amount of precipitation.

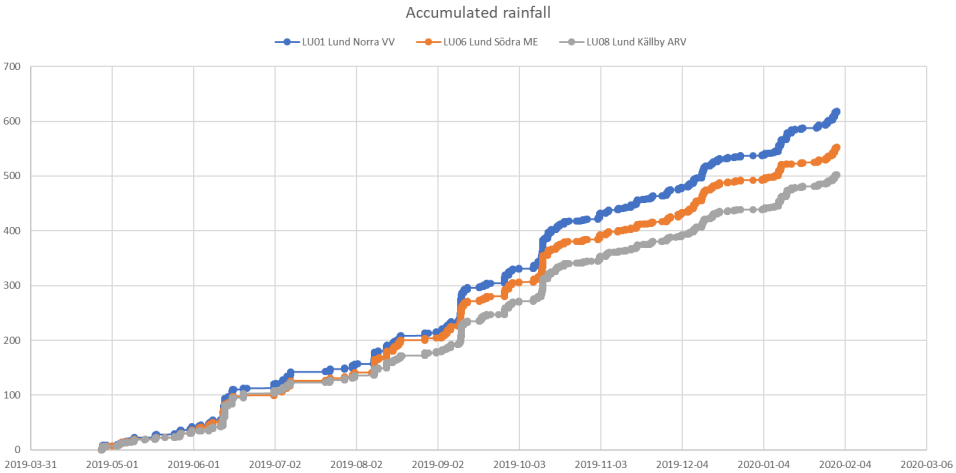

Figure 13: Accumulated rainfall for three rain gauges placed around Lund.

As seen in figure 14, grid points 1 – 5 behave similarly although point 1 records slightly less precipitation than the others. However, all these grid points are considered to be working. Additionally, the grid points show the same pattern as the rain gauges in the previous figure.

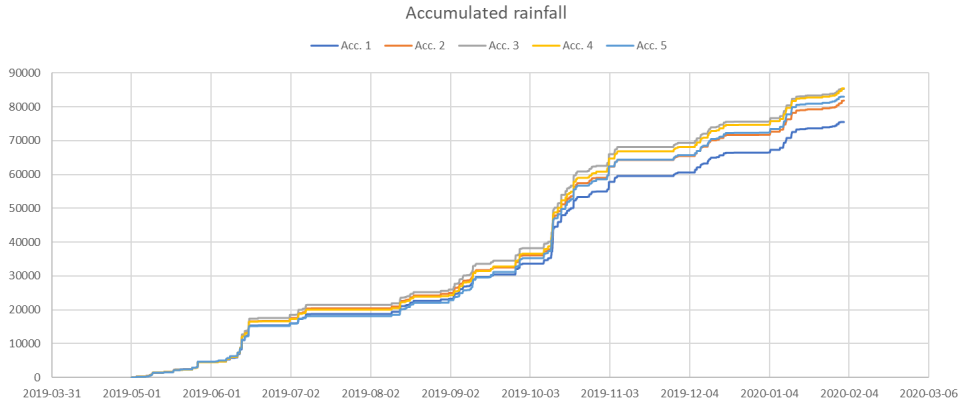

Figure 14: Accumulated rainfall measured by the radar for points 1 – 5.

The same procedure is carried out for the remaining grid points. The figures with accumulated rainfall for these grid points are found in Appendix B – Accumulation of precipitation in grid points.

After investigating the quality of the radar data, it is concluded that the radar indeed suffers from errors. In certain areas, mainly in the southern and eastern parts of the city, the radar seems to either overestimate the amount of precipitation falling or stops recording data completely after the beginning of November. The status of the examined grid points is illustrated in figure 15. The points that did not record any precipitation at all after November are marked in red, the ones that only partly recorded precipitation are marked in yellow, and the unmarked points are considered to be working. The unmarked points might though be slightly overestimating the precipitation falling.

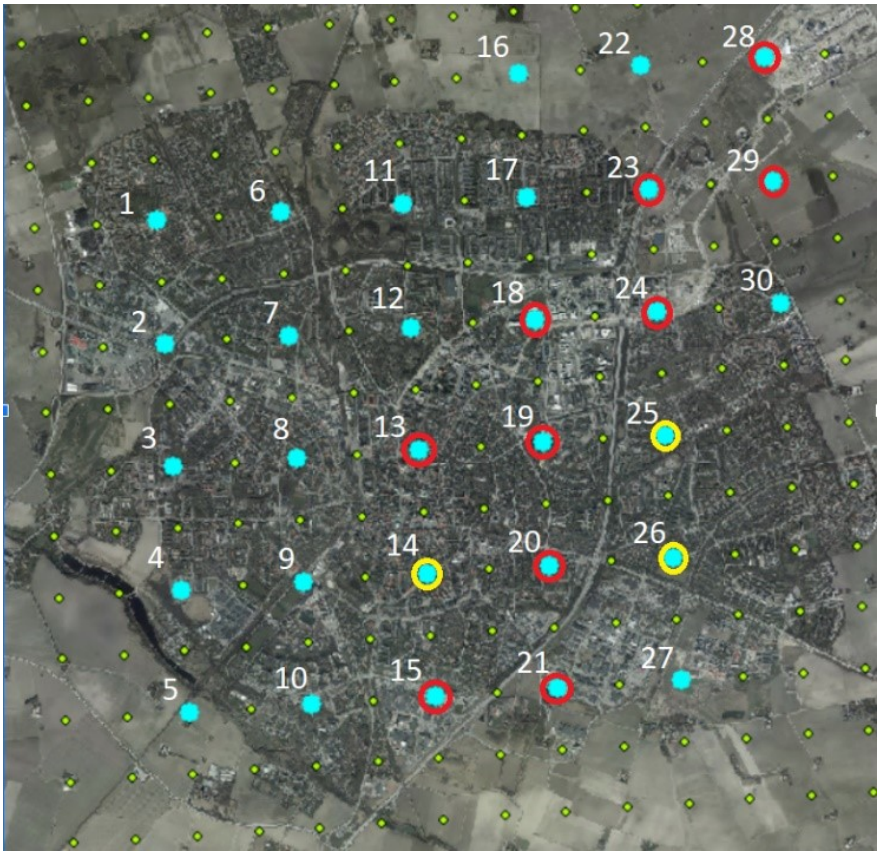

Figure 15: Map of Lund centre. The points marked in red are points where the radar does not measure any precipitation after the beginning of November 2019 and the points in yellow are where the radar only partly measures precipitation for the same period. The green points are equally spaced with 500 meters, spacing the investigated points at 1000 meters.

3.2.4 Attenuation Issues

Additional problems concern attenuation, as described in section 2.3.2. During certain events, where an example is shown in figure 16, special patterns are identified, making the precipitation appear circular around the radar and showing no precipitation outside of this circle. For the example in figure 16, the area with rainfall stays within a constant distance from the radar unit in Dalby and does not move outside of this for a long period of time, which is interpreted as unnatural and likely caused by the attenuation effect. The large areas to the west and southwest of the radar are also not likely areas with no precipitation, but rather areas where the clutter filter has removed the signal altogether. Both rain gauges and the inflow into Källby show that there was heavy rainfall occurring on this day.

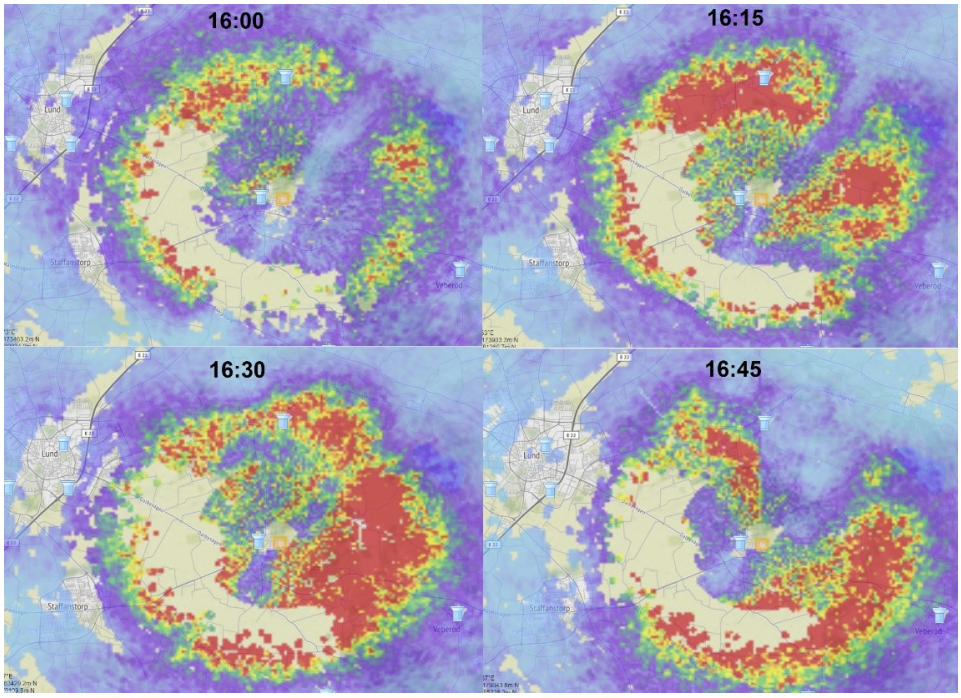

Figure 16: Radar images for the 10th of January 2020 which shows a clear example of attenuation. Dalby is in the centre of the images and the yellow square is the location of the X-band radar unit. The "hole" with missing data over Lund is seen to the left (VA SYD 2020).

3.3 Presentation of Cases

In this section, the cases used in the project are presented. Henceforth, the term *case* is only used in relation to the separate spatial aggregations described in this section.

3.3.1 Division of Catchment and Retrieval of Radar Data

The radar data can be retrieved as either individual grid cells or as an area of aggregated grid cells, referred to as polygons. For this project, three different cases of polygons are defined for the Lund urban catchment seen in figure 5. These are referred to as case 1, 2 and 3. Each case has a different catchment boundary division, in order to investigate what effect this has on usefulness of the radar data used in the neural network. A fourth case, using a rain gauge as source of precipitation input, is also introduced as a reference case.

The hydrological boundary defined by VA SYD is used as outer boundary for the cases, as almost all precipitation that falls outside this boundary is

considered to either not reach Källby WWTP or to have a negligible impact on the flow. The high amount of leakage into the sewage system, further described in section 2.1.3, motivates the use of the entire hydrological area and not just the areas close to the combined sewer system. Smaller areas to the north and northeast of Lund are excluded from the polygon since these areas being assumed to have close to negligible flow contributions.

Although the polygons are defined by the hydrological boundaries, the retrieval of the radar data is restricted to a 500 x 500 m grid. Even if only small parts of the grid cell lie within the polygon, the entire grid cell is included in the aggregated area. This means that in reality, the boundaries of the cases may not match precisely with the hydrological boundaries. There is also a possibility that the sub-catchments in the third case slightly overlap each other because of the grid. Because of this it is important to consider the restriction of the grid, but it is not assumed to have a large impact on the result.

3.3.2 Case 1 - Polygon

In case 1, all radar data within the hydrological boundaries of Lund are used. The polygon is presented in figure 17. The entire area of Lund is used even though the radar data is flawed in certain areas, motivated by that the neural network is clever and can draw conclusions from the input that is fed to it, even if the data is flawed. The benefit of case 1 is its simplicity as no division of sub-catchments is made. If case 1 shows great results compared to the other cases, it indicates that no complicated division of sub-catchments is necessary and also that the model can handle some clutter in the radar data.

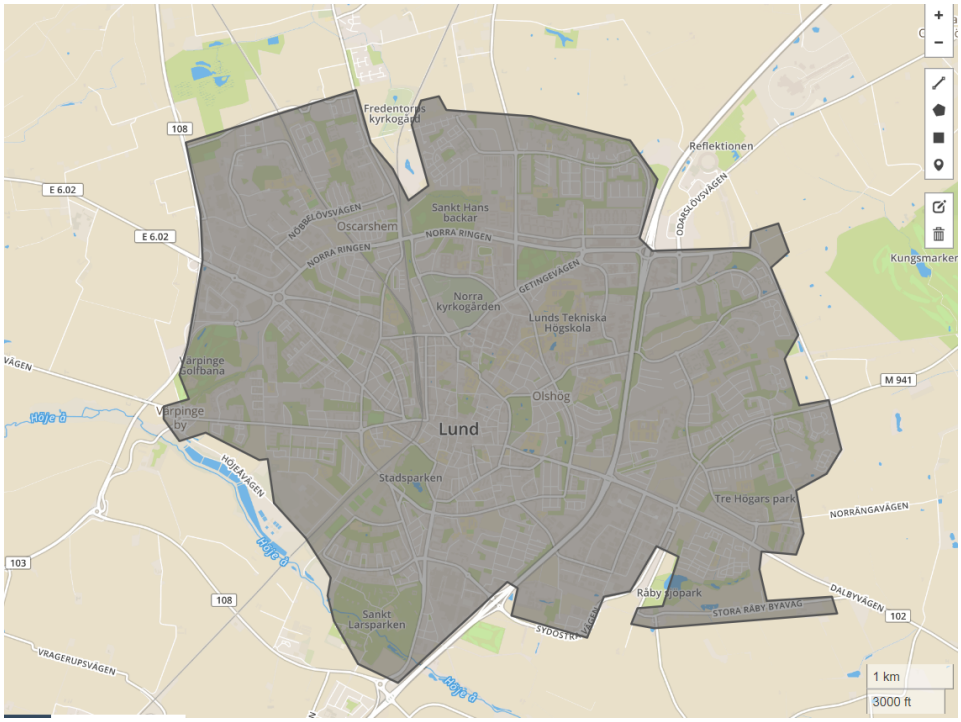

Figure 17: The catchment used for case 1. The catchment corresponds to the hydrological boundaries of Lund city.

3.3.3 Case 2 - Adjusted Polygon

In case 2, only areas where the radar performs well are used. The area is illustrated in figure 18. The same simplicity argument as for case 1 can be used, with the exception that the parts performing poorly are excluded. This is motived by that the neural network might perform poorly if it is fed rubbish. Even though it might seem troublesome to exclude such large areas of the city, it is considered reasonable as these areas have significant defects.

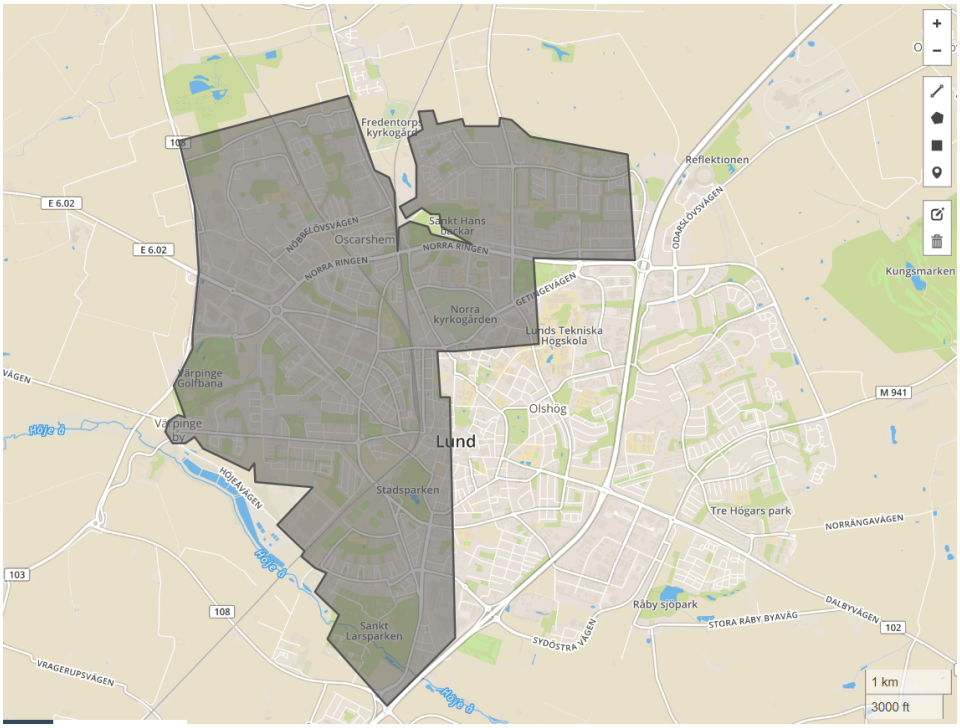

Figure 18: The catchment used for case 2. The catchment corresponds to the hydrological boundaries of Lund with the radar vice non-working parts excluded.

3.3.4 Case 3 - Sub-Catchments

In case 3, the same well-performing area as in case 2 is used. The area is further divided into 4 sub-catchments, north, west, south, and central Lund, see figure 19. The sub-catchments are divided after the hydrological boundaries of Lund, presented in figure 5. Using several sub catchments is thought to aid the network in making forecasts. Certain areas might have general characteristics, such as large amounts of impermeable surfaces or greenery, which could be useful information for the network. Additionally, it could be argued reasonable to include an eastern sub-catchment with the defect radar data, as the area might contain useful information, even though the data is flawed. However, this is not considered to be the case and no eastern sub-catchment is included.

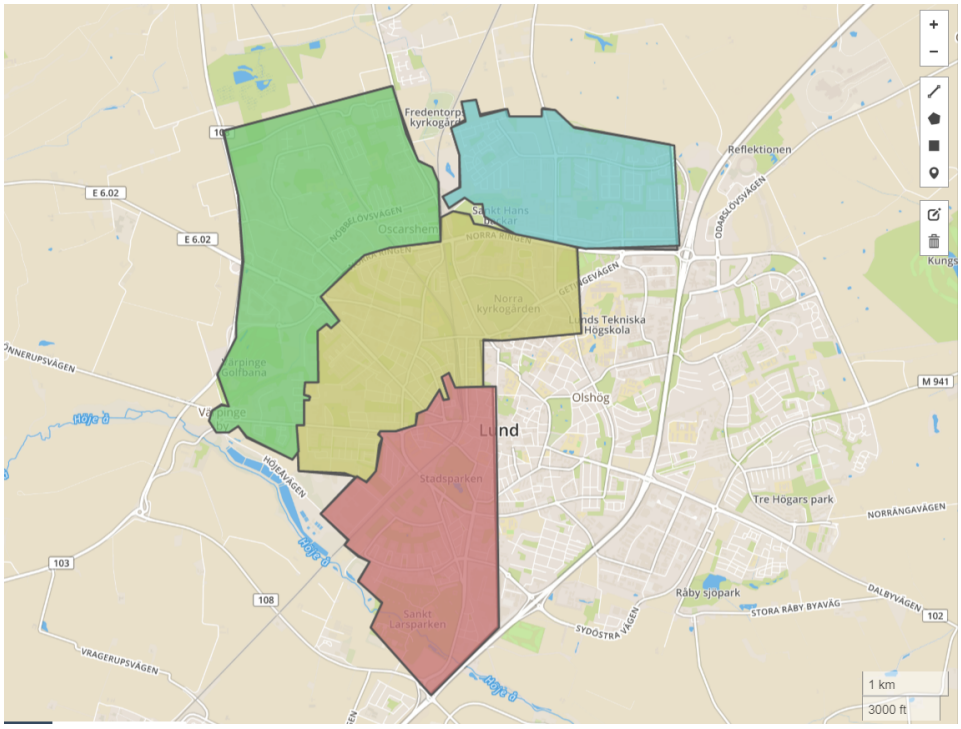

Figure 19: The sub-catchments used in case 3. Green represents the western area, blue the northern area, red the southern area and yellow the central area.

3.3.5 Case 4 – Reference Case

To evaluate and compare the results of the 3 other cases, a fourth reference case is introduced. The fourth case is not using radar data but rain gauge data as precipitation input, as the conventional way of measuring precipitation to a large extent is and previously has been done by using rain

gauges.

The rain gauge used is labelled LU08, which is located at Källby WWTP, see figure 4. The accumulated rainfall for LU08 is illustrated in figure 13.

3.4 Input Variables

As previously stated, one or if not the most crucial part of the design process related to the neural network is the selection of input variables, as the network will struggle with making a forecast if it is handed bad input data. Therefore, the neural network will be tested with several different types of input variables in order to evaluate the significance of each one for the task of forecasting inflow into Källby WWTP. Table 2 shows the studied variables. These input variables are selected as they are thought to provide relevant information about the catchment or sewer system.

Table 2: The input variables studied.

Function	**Data Type**	**Source**	**Resolution**
Main source of precipitation data	Radar data	VA SYD	1 min
Additional wastewater information	Dalby wastewater flow	VA SYD	Approx. 25 min
Hydrological conditions in catchment	Höje å discharge	SMHI	1 d
Hydrological conditions in catchment	Groundwater level	Trafikverket	5-25 d, avg. 10 d
Characteristics of precipitation	Wind information	SMHI	60 min
Additional source of precipitation data	Rain gauge Källby WWTP	VA SYD	1 min
Target and current WWTP inflow	Källby Inflow	VA SYD	3 min

3.4.1 Radar Data

The primary input variable is the X-band radar data. The X-band radar data is obtained as total volume of rain with the temporal resolution of 1 min for either a 500 x 500 m grid cell or a manually defined spatial polygon. The accumulated total volume of rain for each time stamp is also based on the size of the catchment and is obtained in cubic meters for every time stamp. The data that is used is raw and non-bias-corrected. All radar data is retrieved from level 3 for the radar, which corresponds to 8 degrees which is roughly 1400 m above Lund (South *et al*, 2019).

3.4.2 Wastewater Flow from Dalby

The wastewater flow from Dalby and surrounding villages, including Veberöd and Genarp, is measured in litres per second roughly every 30 min by VA Syd. Since parts of Dalby have a combined system, the wastewater flow also contains stormwater. This means that in case of rain events, the flow from Dalby will greatly increase, potentially in a similar manner to the flow measured at Källby. This variable is referred to as *Dalby* in the testing because the combined wastewater flow from surrounding villages is

measured in a pumping station in Dalby.

3.4.3 Hydrological Conditions in the Catchment

The X-band radar measures the precipitation in the air instantaneously, but in order to simulate a memory of what precipitation that previously has fallen as well as seasonal patterns, the groundwater level and discharge in the watercourse Höje å are tested as input variables. Höje å is the main recipient in the catchment and flows right past by Källby WWTP, and the discharge may give a good indication of the general hydrological conditions. Data of the discharge in Höje å is provided by SMHI. The variable with discharge in Höje å is referred to as *Höje å* in the testing.

Groundwater data is provided from the Swedish Transport Administration (Trafikverket), which measured the groundwater levels during track work on the train line between Lund and Arlöv during the period May 2019 to January 2020. The average value from 21 measurements of the groundwater level in proximity to Källby WWTP is used. As only the variation in groundwater level is considered useful information, the average value likely gives a realistic pattern of how it varies over the year. The groundwater level is measured in meters above sea level.

3.4.4 Wind Direction and Wind Speed

The wind direction and wind speed in Malmö and Hörby are tested to see if there is any correlation between these variables and the flow at Källby WWTP. Malmö and Hörby are selected as there currently is no wind speed and wind direction data available for Lund. A study by Johansson and Chen (2003) shows that there is a relation between the amount of precipitation falling and wind speed, as well as a relation between wind direction and precipitation patterns. The wind speed and wind direction are measured by SMHI and is updated every hour.

3.4.5 Rain Gauges

Rain gauges are tested as an input to see if they can complement the radar with information. As mentioned previously, one radar-related issue is attenuation where the radar signal gets blocked by water particles in case of an intense rainfall, making it unable to register precipitation that falls behind the curtain of rain. The gauges may assist the network with information that the radar missed. The rain gauges are of a tipping bucket model, registering a value for every 0.2 mm precipitation that falls. The rain gauges are operated and maintained by VA SYD.

Because of technical issues related to the treatment of the rain gauge data, it was not possible to register more than one value per minute in the rain gauge time series. This has a negative effect on the data from the larger precipitation events, as these are the times when multiple values might register each minute. However, as the rain gauges are not part of the main aim of the study and the data loss is minor, this issue is not considered to be of major importance.

3.4.6 Inflow into Källby WWTP

The flow into Källby WWTP is measured every third minute in the unit litres per second. The inflow shows a clear diurnal pattern, as seen in figure 20, which is common for wastewater flows. The figure has uncalibrated precipitation data from the south polygon of case 3 added as a visualisation of the response to rain. The flow normally varies between approximately 200 to 500 l/s every day but usually increases after rain events. The increase varies but any flow above 600 l/s is most likely a result of rain. Analysis of the time series shows that roughly 95 % of the inflow data has values smaller than 600 l/s. The data with inflow to Källby is referred to as *Target* when used for comparison with forecasts made by the model, and simply as *Källby* when included as a variable in a test.

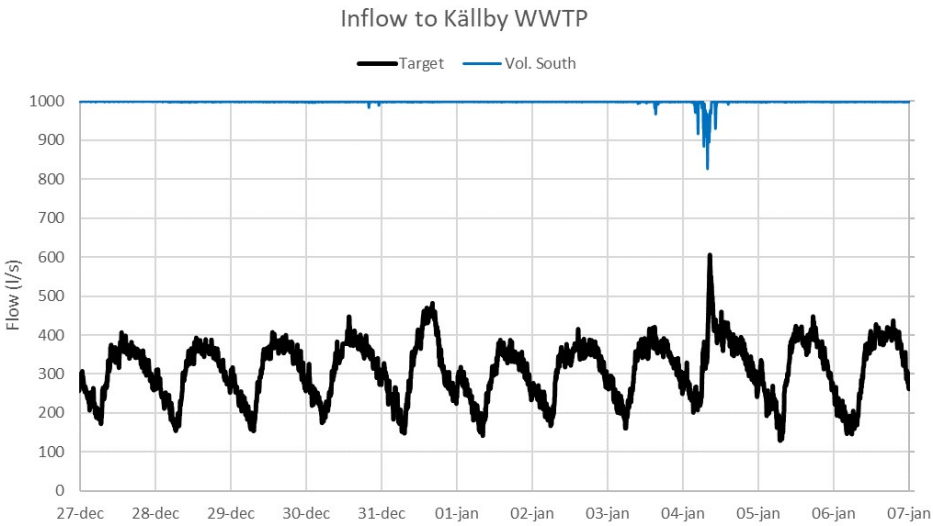

Figure 20: Principal graph showing both the inflow into Källby WWTP and how it normally reacts to a rain event. Vol. South is radar data for the south polygon of case 3.

The values in the time series for the inflow to Källby never exceed roughly 2 200 l/s. This is likely a technical issue in the flow gauge where it cannot

measure flows over this limit. On a few occasions during the study period the flow appears to be restricted by this when in reality it most likely exceeded 2 200 l/s. There is also data missing on a few other occasions, ranging from minutes to hours at a time, which are left missing.

3.5 Evaluation of Performance

3.5.1 Evaluation Method

To evaluate the performance of the network, several days where larger flows were observed at Källby are studied. These days are presented in table 3. A large flow is defined as a flow larger than 800 l/s and days with large flows are referred to as *events*. Several parameters with the purpose to indicate the general performance of the model are also evaluated. Flows that increase a lot in a sudden manner are flows that are most crucial for the wastewater plant to be able to forecast. Therefore, a good forecast of the flow is defined as a prediction that can capture the timing and value of a sudden increase in flow.

Table 3: The events studied.

Date	**Max. flow Källby (l/s)**	**Duration**	**Acc. precipitation (mm)**	**Number of Peaks**
13/08	1 100	2 h	10.2	1
27/09	1 850	4 h	11.6	1
11/10	2 200	3 h	31.6	1
02/11	1 000	3 h	7.6	1
06/12	900	2 h	3.8	1
15/12	900	3 h	6.2	1
10/01	1 100	10 h	9.2	2
12/01	1 200	10 h	11.2	1

For the entire time series, the root mean square error (RMSE), R^2 and loss are evaluated. The root mean square error is calculated with equation 1

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(prediction_i - observed_i)^2}{n}} \tag{1}$$

where n is the number of observations. The formula is applied to all data points of each simulation. The RMSE is based on the residuals, which in this case is the difference between observed and predicted forecasted value for the same time stamp (Barnston 1992). The RMSE for an entire test period will be a value larger than 0 and can be used for comparison between tests with the same type of data. A smaller value on RMSE indicates better fit.

Plotting all predicted values against the corresponding observed value for

each time stamp in the time series, the values would ideally form a 1:1 relationship. A real linear regression for the model is optimised and plotted in a graphing software. Should the model generally over- or underestimate, this linear regression would be steeper or flatter than the 1:1 relationship. For the linear regression calculated, the R^2 value is also computed, which gives an indication on the general deviation from the data points from the regression equation. It is calculated with equation 2.

$$R^2 = 1 - \frac{\sum(observed_i - prediction_i)^2}{\sum(observed_i - mean_{obs})^2} \tag{2}$$

The loss for the entire validation period is also used as an indication of the error in each model. The value of the loss is obtained from Tensorflow and further information about loss is found in section 2.4.3. The R^2, RMSE and loss values are referred to as the *evaluation parameters*.

The evaluation parameters describe the model's general ability to make a forecast and are good indicators of how well the model performs. However, they might not necessarily describe the model's ability to predict the timing and value of peaks. When there are no large precipitation events, which is the case for most of the time series, the wastewater flow follows a daily and weekly pattern. The model might be good at estimating these patterns, yielding high scores from the evaluation parameters. Therefore, it is important to not blindly trust these parameters, as what is important is the peak performance. Instead, the performance of the model should be concluded from both the value of the evaluation parameters and the physical appearance of the graph.

The input variables that are tested individually and together to conclude how a certain mixture of variables affect the result, can be seen in table 4. Additionally, sine functions and roll durations as described in section 3.1 are included in all tests. These are not considered as input variables in this context but rather as variables that are constantly present in all tests.

Table 4: The different tests with associated input variables. Radar means by radar measured volume per minute and rain gauge refers to the rain gauge at Källby treatment plant (LU08).

Test nr.	**Input variables**
0	Radar
1	Radar, Dalby
2	Radar, Höje å
3	Radar, Dalby, Höje å
4	Radar, rain gauge
5	Radar, rain gauge, Dalby, Höje å
6	Radar, wind information
7	Radar, groundwater level
8	Radar, Källby WWTP
9	Radar, Källby WWTP, Dalby, Höje å
10	Radar, rain gauge, Källby WWTP, Dalby, Höje å

3.5.2 Prediction Time

It is of interest to investigate how far ahead in time a reliable forecast of the flow can be made with the model. Therefore, the prediction time of the best performing test is evaluated.

Various prediction times are tested when investigating how the prediction time affects the result. The prediction times tested are in 15-min steps from 45 min to 120 min, as visual inspection of the data indicates that there seldom is much longer time between precipitation and increase in inflow at Källby WWTP.

3.5.3 Further Extension of the Prediction Time

In this study, all precipitation input information is connected directly to Lund, which will limit how far ahead a forecast of the flow can be made. As previously stated, is there seldom, if ever, more than 120 min between observed precipitation and increase in inflow to Källby. To further extend the forecast other input information is needed, which potentially could be in form of a forecast of the precipitation.

The ability to forecast flows up to 4 h ahead of time would be of great value for the wastewater treatment plant as well as for VA Syd's other operations in general, which would provide plenty of time to prepare for potential large flows and floodings. Therefore, it would be of interest to investigate if it hypothetically possible to, with an accurate forecast, produce lengthy

forecasts of the flow on this scale with the neural network.

To imitate a perfect forecast of precipitation, the time series of a rain gauge in Lund is advanced 4 h ahead in time and used as precipitation input information. The prediction time is set to 4 h and the model additionally uses information from Dalby and Höje å as inputs. The purpose of the 4 h forecast is to evaluate the potential of long forecasts made by the neural network, by using hypothetical data.

3.5.4 Comparison to Conventional Model

The quality of the best 60 min forecast made by the neural network in this thesis is compared to that of a simulation of the same inflow time series made by a conventional MIKE URBAN model used at VA SYD. The conventional model uses rain gauges as input data. The comparisons are evaluated by visually comparing the performances of the different models during individual rain events. The purpose of the comparison is to see how the forecast of the neural network compares with the simulation of a conventional model, as this is what is available for comparison.

It is important to distinguish between the different functions and purposes of the models. The neural network in this thesis is used to forecast the flow into Källby 60 min into the future while the conventional model is used to simulate how the sewer system in Lund reacts to rain events. These differences need to be considered when comparing the two models. Additionally, the purpose of the conventional model is not to forecast a flow, but rather to simulate how the flows distribute throughout the system as a reaction to rain events which makes it a useful planning tool. The rain time series used in a conventional model can either be collected data or hypothetical artificial rain time series.

4 Results

In this section, the results are presented. Initially, case 1, 2 and 3 are compared regarding loss, R^2 and RMSE in order to determine which case should be used for the testing of variables. The tests are found in table 4. First, the tests are compared. Then, a selected few are optimised with regard to model parameters. After this, the selected tests are further evaluated regarding error and uncertainty, followed by an investigation of the prediction time. Lastly, the forecast of the best performing test is compared with the simulation of a conventional model.

4.1 Results from Analysis of Cases

In this section, case and test names are shortened, meaning that for example case 1 test 3 can be referred to as either case 1.3 or case 1_3.

4.1.1 Evaluation Parameters

To determine the overall performance of the three cases, the cases are compared with each other regarding loss in the training period, RMSE and R^2, illustrated in figure 21, figure 22 and figure 23. Case 4 acts as a reference case when comparing the model performances but is not considered when selecting the best performing case. Case 4.4, 4.5 and 4.10 do not exist as they would be replicas containing the same inputs as other existing tests.

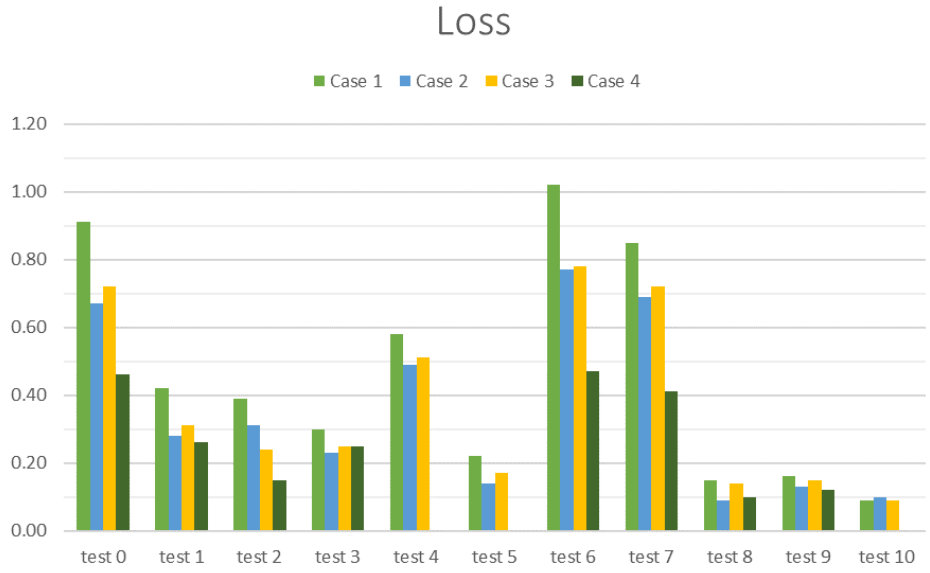

Figure 21: The loss presented for each case for each test.

As seen in figure 21, case 4 has the lowest loss for all tests, which indicates that the network is better optimised using rain gauge data as opposed to radar data. Case 1 continuously has the highest training loss, indicating that the model has a harder time finding the optimal solution for this case. Test 10 is the test with the lowest calculated loss. Case 2 and 3 perform similarly for all tests. Most clear from this analysis is perhaps that tests 0, 6 and 7 have considerably higher loss than most other tests. These tests include wind and groundwater data, which implies that these sources of information do little to improve the effectiveness of the network.

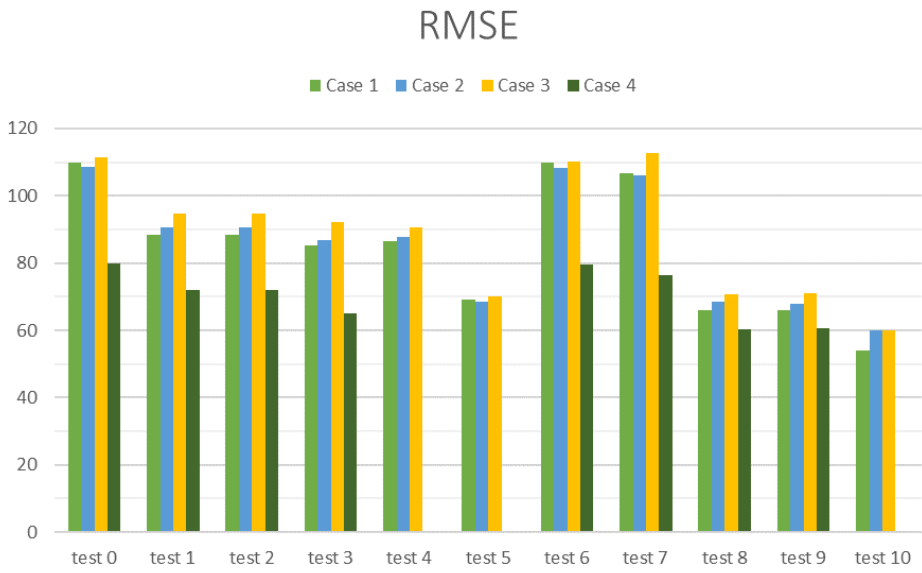

Figure 22: The RMSE presented for each case for each test.

Figure 22 shows how the RMSE value barely changes between cases 1, 2 and 3, but is consistently lower for case 4. This indicates that the different catchment definitions are not of significant importance for the RMSE. Excluding case 4, case 3 performs slightly better than the other cases. Test 10 is the test with lowest measured RMSE.

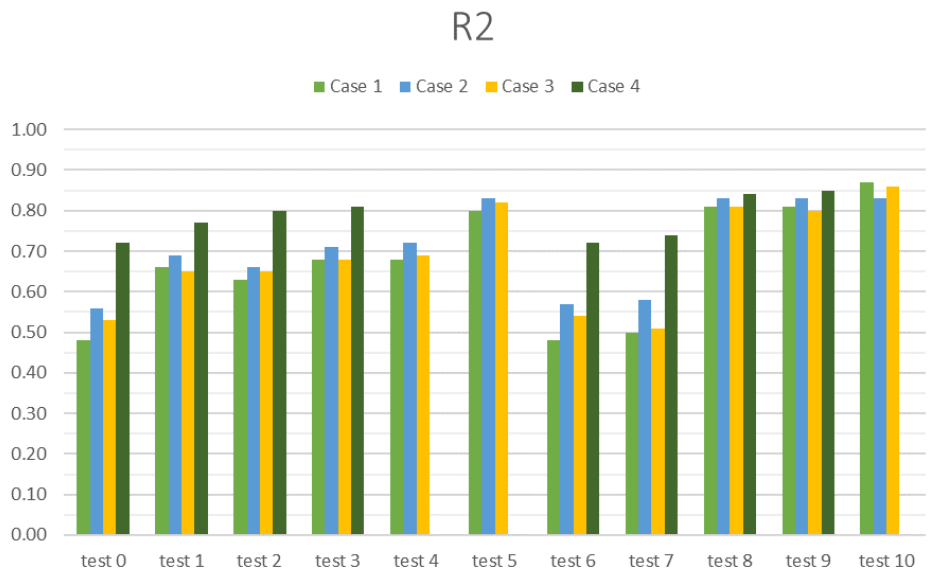

Figure 23: The R^2-value for each case for each test.

The R^2 value is presented in figure 23. Case 4 outperforms all other cases in test $0 - 7$ and performs just slightly better than the others in test 8 and 9. Test 10 is also the best performing test with regard to this parameter. Case 2 performs slightly better than case 1 and 3 for most of the tests.

Excluding the reference case, case 2 and 3 are concluded to be the overall best performing cases overall, with regard to loss, RMSE and R^2. As case 1 by these evaluations does not perform as well, it is excluded from further comparisons and selection of the best performing case.

Studying case 2 and case 3 visually for individual rain events, the cases seem to perform similarly but case 3 generally overestimates the flow to a larger extent than case 2, but with a faster response. A visual comparison can be seen in figure 24. The black curve represents the measured flow at Källby WWTP and the blue and red curves are the 60 min forecasts of cases 2 and 3 respectively. On the top of the graph, the precipitation measured by the radar is visualised. As the radar data is uncalibrated, the precipitation is only used to visualise the timing of the rain.

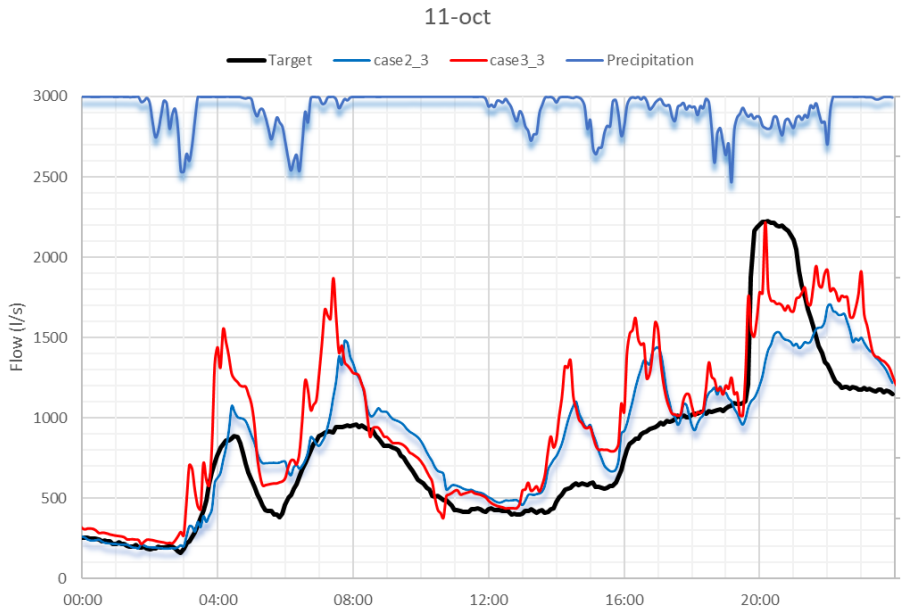

Figure 24: The response of test 3 for case 2 and case 3 during the rain event of 11 October 2019.

Judging from this and the evaluation parameters, case 3 is considered the best performing case. The evaluation parameters showed no significant advantage to either case, only a slight advantage to case 2. However, the quicker response to precipitation of case 3 is highly valued, thus making case 3 the best performing case. The similar performance of the cases indicates that further division of the catchment into sub-catchments is not of great importance for the model.

As case 3 provides more spatial information and responds faster to precipitation, it is selected the best performing case. Therefore, case 3 will henceforth be used when evaluating the input variables.

4.2 Analysis of Tests and Input Variables

When evaluating the input variables, the different tests are compared with test 0 which serves as a reference test as it only uses radar data. In this way the impact of the different input variables can be evaluated relative to using only radar data. The graphs presented are considered to be representative for the behaviour of the different tests.

4.2.1 Test 0, 1, 2 and 3

Test 0 (radar data) for case 3 delivers results according to figure 25. It responds quickly to precipitation and recognises that there should be a peak, even though the value of the peak is slightly inaccurate. However, test 0 sometimes struggles to estimate the base flow when it is not raining, indicating that the model needs additional information to improve this. Clear examples of this can be seen in figure 27 and figure 28.

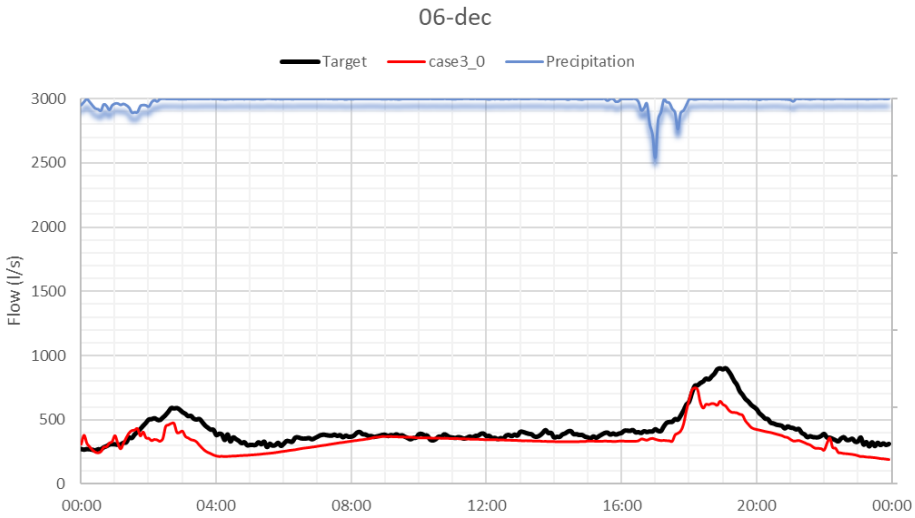

Figure 25: Observed and predicted flow at Källby WWTP the 6th of December 2019.

According to the evaluation parameters in figure 21 to figure 23, both the information from Dalby and Höje å seem to increase the performance of the model and combining them both further increased the performance. By studying the behaviour of the tests in figure 26, it is seen that both test 1 (radar and Dalby) and 2 (radar and Höje å) overestimate the flow compared to test 0. However, test 1 and 3 (radar, Dalby and Höje å) react quicker to the precipitation than test 0. This further indicates that information about the hydrological conditions in the catchment and additional flow information from the sewage system indeed improve the model.

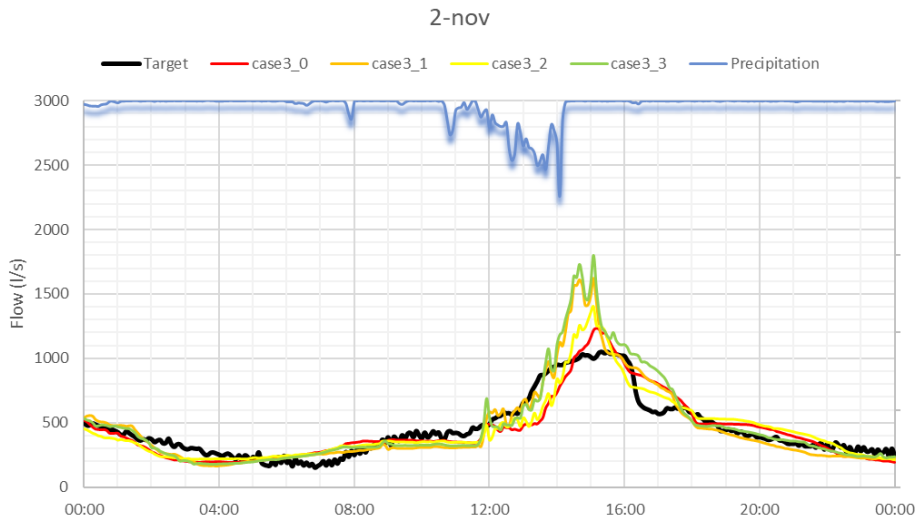

Figure 26: Observed and predicted flow at Källby WWTP the second of November 2019.

Figure 27 indicates that test 1 and 2 provides information to the model where the radar data alone is not enough to make a good forecast. Test 1 decently imitates the shape of the second peak while test 2 is next to parallel with test 0. Test 3 performs slightly better than both test 1 and 2, making a good forecast even though the radar data is sparse.

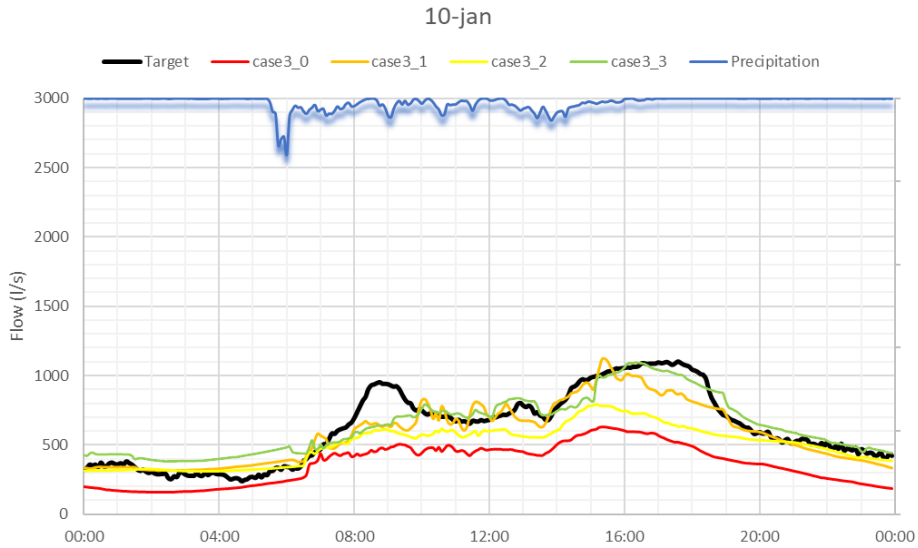

Figure 27: Observed and predicted flow at Källby WWTP the 10th of January 2020.

Comparing 10th of January with other days, the flow at Källby WWTP is very large compared to the amount of precipitation registered by the radar. This could be due to various reasons, but likely due to attenuation issues. Adding additional information together with the radar data greatly improves the forecast.

For the 6th of December in figure 28, test 1, 2 and 3 perform similarly. Test 2 struggles a bit with the peak at 18:00 but gives a good forecast of the base flow.

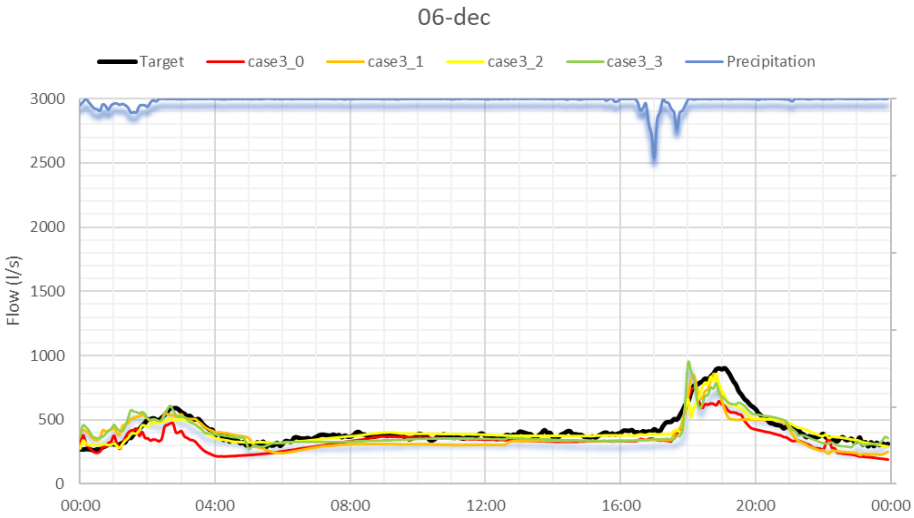

Figure 28: Observed and predicted flow at Källby WWTP the 6th of December 2019.

11th of October was a day with large measured flows at Källby WWTP. All tests struggle with making a good forecast of the flow, as they consistently overestimate the peaks throughout the day but fail to properly estimate the large peak at 20:00, see figure 29. However, test 3 is the test that most accurately captures the large peak at 20:00. The actual peak could potentially have been even higher, as it is not possible to properly measure flows larger than roughly 2 200 l/s at Källby WWTP.

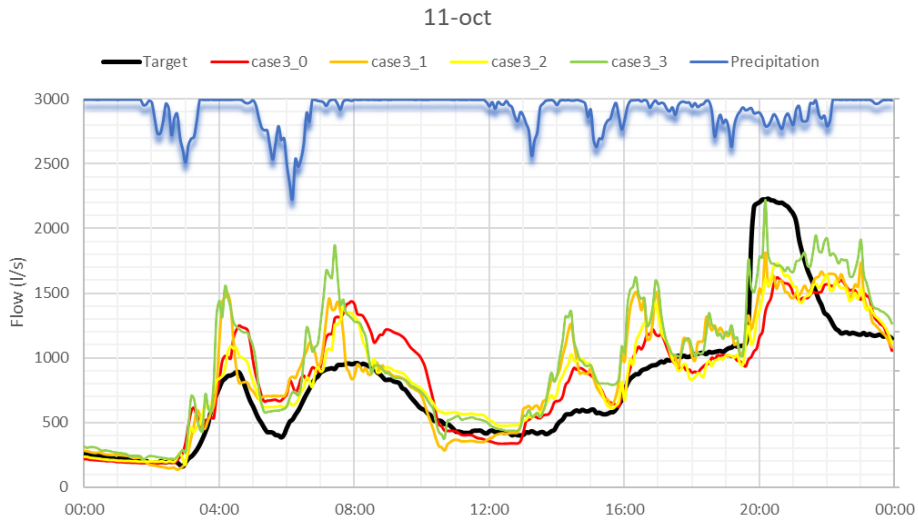

Figure 29: Observed and predicted flow at Källby WWTP the 11th of October 2019.

4.2.2 Test 4 and 5

Adding rain gauge data aids the model in making forecasts when the radar data suffers from issues, for example attenuation, see figure 30. However, the rain gauges do not improve the forecast of the base flow. By additionally adding information from Dalby and Höje å, the model's ability to forecast base flow is improved. Overall, the rain gauge supplies the radar data-based model with valuable additional information.

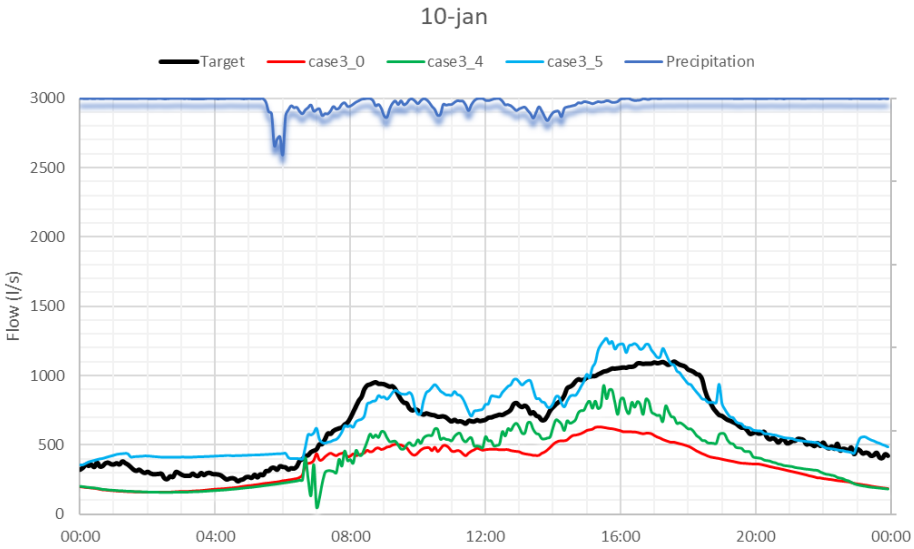

Figure 30: Observed and predicted flow at Källby WWTP the 10th of January 2020.

As seen in figure 31, test 4 (radar and rain gauge) and test 5 (radar, rain gauge, Dalby and Höje å) show similar behaviours, with the difference that test 5 is shifted roughly 400 l/s higher because of the difference in base flow. Test 5 performs better than the other tests, likely because of the additional flow information from Dalby and Höje å.

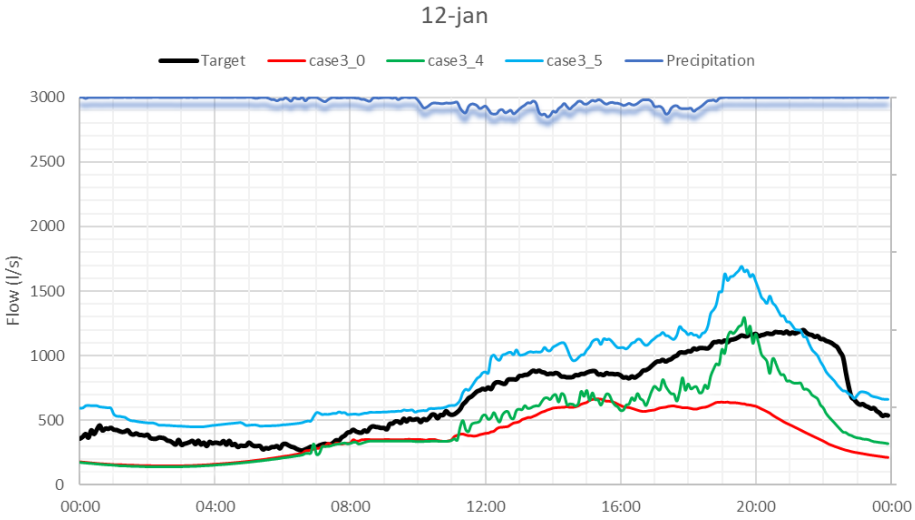

Figure 31: Observed and predicted flow at Källby WWTP the 12th of January 2020.

As seen in figure 32, all tests follow a similar trend on the 11th of October. However, the rain gauge information helps test 4 and 5 to predict the high flows at 20:00, which the model is not able to do properly without the information from the rain gauge. This is likely because of attenuation effects affecting the radar.

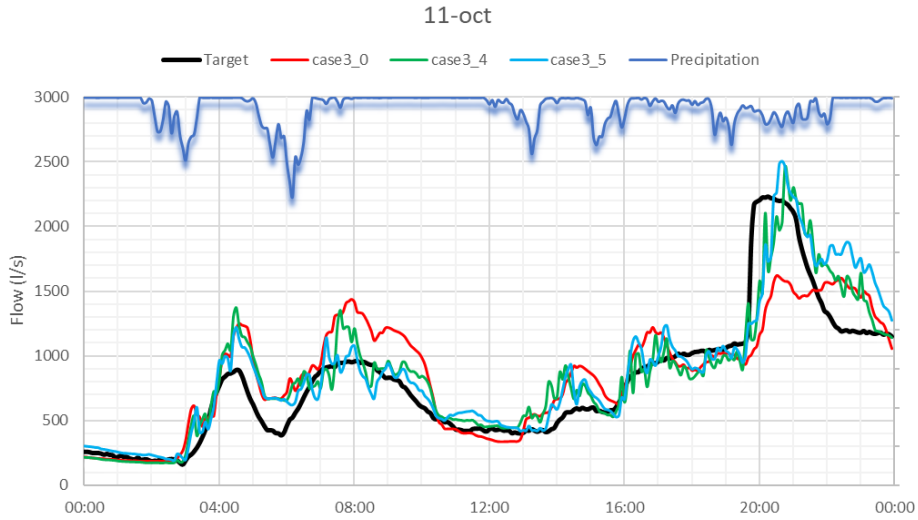

Figure 32: Observed and predicted flow at Källby WWTP the 11th of October 2019.

4.2.3 Test 6 and 7

By studying the evaluation parameters in section 4.1.1, there is no difference in the performance of test 0 (only radar), test 6 (radar and wind) and test 7 (radar and groundwater), indicating that neither wind nor groundwater information does not improve the performance of the model. As seen in figure 22, there are only minimal difference between the tests. The base flow behaves the same and the peak performance is similar, additionally indicating that these variables are useless to the model.

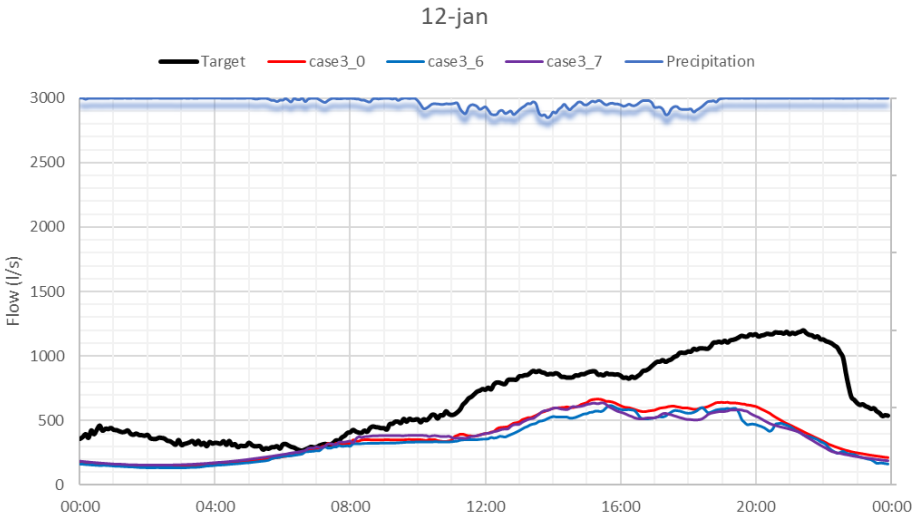

Figure 33: Observed and predicted flow at Källby WWTP the 12th of January 2020.

Test 6 and 7 follows test 0 almost identically for most of the time period. Test 7 is sometimes more responsive to precipitation than test 0, illustrated in figure 34. However, this is not a consistent behaviour.

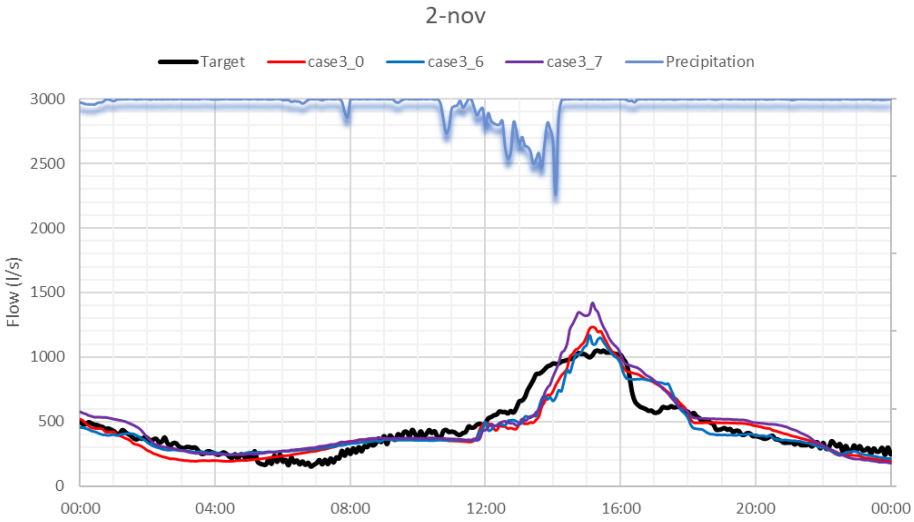

Figure 34: Observed and predicted flow at Källby WWTP the 2nd of November 2019.

4.2.4 Test 8, 9 and 10

From the result of the evaluation parameters it is clear that information about the flow at Källby WWTP improves the model. This is confirmed by studying the events, which illustrates that the model becomes more accurate in its forecast. In figure 35, up until 12:00, test 8 (radar and Källby), 9 (radar, Källby, Dalby and Höje å) and 10 (radar, rain gauge, Källby, Dalby and Höje å) tend to a lesser extent overestimate the flow compared to test 0. However, the rain gauge information in test 10 shows to be valuable as test 10 best predicts the peak at 20:00, while the other tests fail to do so.

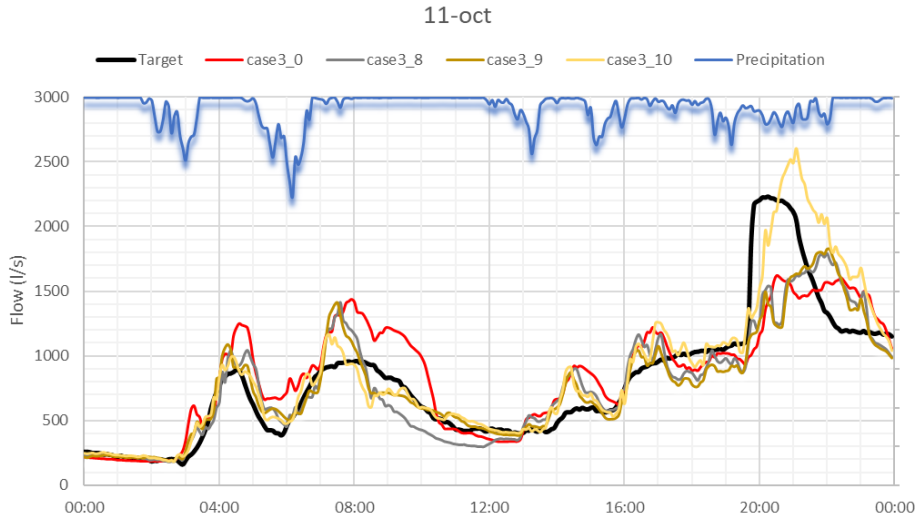

Figure 35: Observed and predicted flow at Källby WWTP the 11th of October 2019.

Figure 36 shows that flow information from Källby WWTP greatly improves the model's ability to forecast the base flow. This is not surprising, as the base flow at the treatment plant normally does not change drastically from one hour to another if there is no precipitation event.

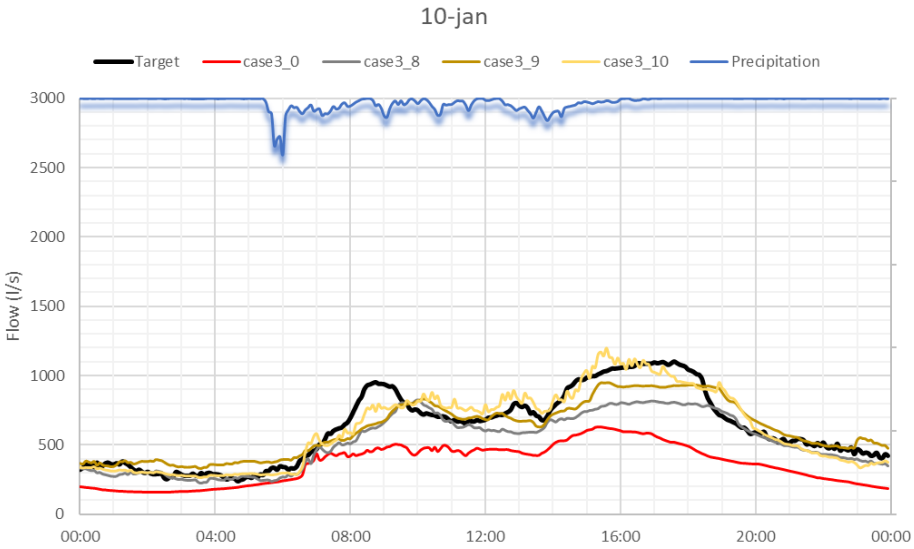

Figure 36: Observed and predicted flow at Källby WWTP the 10th of January 2020.

4.2.5 Conclusion of Tests

In table 5 the improvements and deteriorations for each test is concluded. The tests are described in table 4.

Table 5: Improvements and deteriorations for each test compared with test 0 (only radar data).

Test	Improvements	Deteriorations
1	- Quicker response to peaks - Events with insufficient radar information/attenuation - Base flow	- Occasionally overestimating flow
2	- Base flow	- Occasionally overestimating flow
3	- Quicker response to peaks - Duration of peaks - Events with insufficient radar information/attenuation - Base flow	- Occasionally overestimating flow
4	- Events with insufficient radar information/attenuation - Overall fit	-
5	- Events with insufficient radar information/attenuation - Base flow - Overall fit	- Occasionally overestimating flow
6	-	-
7	- Sometimes more responsive to precipitation	-
8	- Base flow - Overall fit	-
9	- Base flow - Overall fit	-
10	- Base flow - Peaks - Overall fit	-

It is concluded that for the tests without rain gauge input information, test 9 is the best performing test. Also considering the tests with the rain gauges, test 10 performs best. This indicates that the model greatly benefits from information about the wastewater system (Källby and Dalby), discharge in catchment recipient (Höje å) as well as nearby rain gauges together with the radar data. The wind information did not benefit the model in itself but would perhaps do so if additional precipitation data outside of Lund was added in connection to the wind gauge locations. This could theoretically give useful information if rain clouds were approaching Lund before any rain had fallen within the catchment.

The groundwater information did not benefit the model either, which likely is because the movement of groundwater is a too slow process to be of use to the model. This indicates that the network benefits more from information about faster processes, or short-term memory of the catchment, such as information from Höje å river.

4.2.6 Comparison Between Best Performing Cases

A comparison between the test using only radar data (test 0), the best performing test without rain gauge information (test 9) and the overall best performing test (test 10) is made with the purpose to visualise how they compare. Both test 3 and test 9 perform decently, but test 9 is chosen as it does not overestimate the flow to such an extent as test 3, although test 3 catches the peak seen in figure 29 better than test 9.

In figure 37 it can be observed that test 10 is the test that generally best forecasts the flow for the 11th of October. Test 0, which only uses radar data, struggles to forecast the peak in the evening and overestimates the flow during the day.

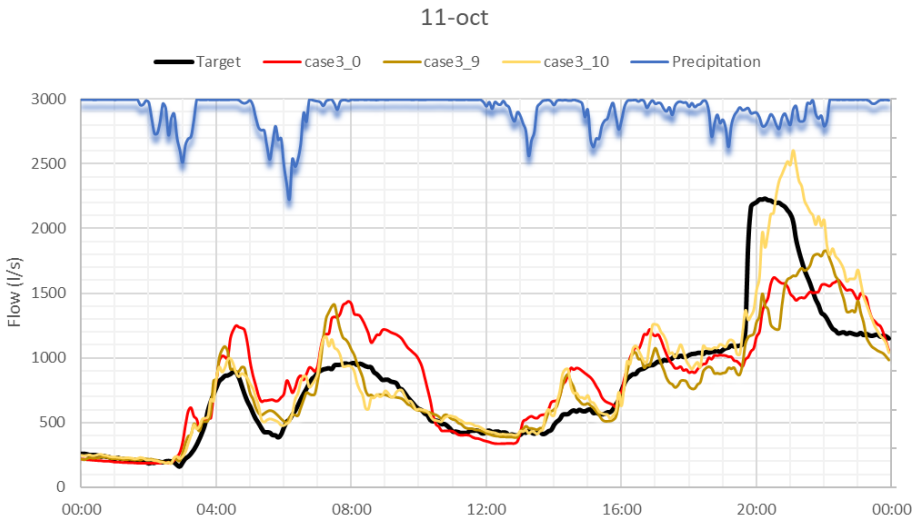

Figure 37: Predicted mean flows from tests 0, 9 and 10 on October 10th.

It can be interpreted from figure 38 that the model sometimes struggles to forecast the flow with only information from the radar data. Comparing test 0 to tests 9 and 10, there is a rather significant improvement in forecasting the flow in the latter two.

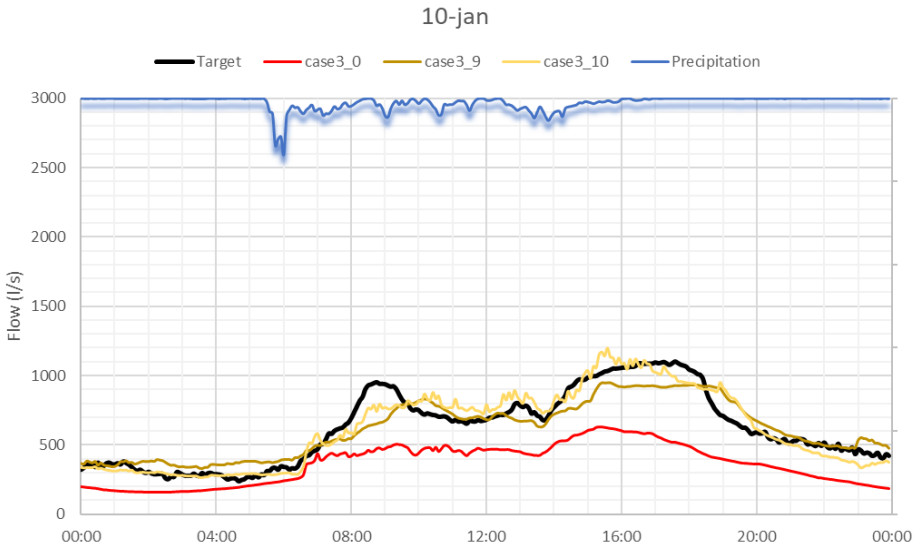

Figure 38: Predicted mean flows for tests 0, 9 and 10 on January 10th.

Figure 39 shows a direct comparison of the forecasted flow and the measured

target flow at Källby WWTP for every timestamp in the whole time period used. As seen in the figure, the vast majority of the data points have values of less than 500 l/s. As indicated by the line pattern on the right side of the figure, there is a limit for how large flows Källby WWTP can measure, which seems to be roughly 2 200 l/s. At this limit there is also a range of forecasted flows, according to figure 39.

The reason for this range of flows could be the fact that the highest peaks often are slightly delayed in the forecast, and as seen in the 20:00 peak in figure 37, where the forecasted flow still is rising while the measured flow is steady at the maximum. It is assumed that if the flow gauge would register all high flows accurately, the line pattern at 2 200 l/s would have more of a cloud shape reaching to the right side of the figure.

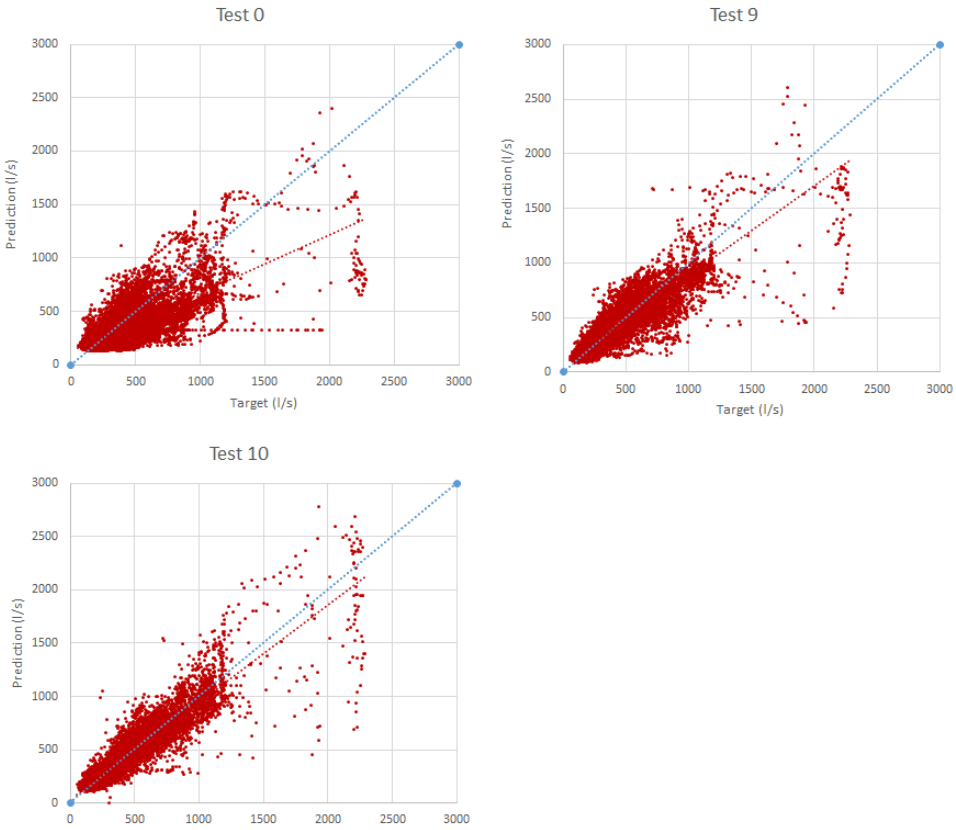

Figure 39: Plots of predicted and measured values above 500 l/s for Case 3 test 0, 9 and 10. The blue line is a visualisation of a 1:1 relationship and the red line is a linear trendline.

By further studying figure 39, there seems to be a threshold at 1 200 l/s for

the measured flow. The reason for this threshold is not known but could be related to operations at the treatment plant. Potentially, there could be certain procedures when the flow reaches a certain level, to keep the operations running smoothly. This level could then possibly correspond to 1 200 l/s.

All tests generally underestimate the flow, but test 10 to a lesser extent than test 0 and test 9 in between. The trend line for test 10 is rather close to 1:1 but there is still some underestimation. The width of the cloud of points indicates the general variation in error for all data points. The fairly narrow cloud that test 10 shows, suggests that the forecasted values do not deviate from the respective measured values as much as the other tests. Further, the scatter plots show no information regarding timing of the peaks, but the large presence of underestimations, including in test 10, indicates that the forecasts of the largest flows are regularly inaccurate. However, as the model is trained on data that never exceeds 2 200 l/s, when it in reality likely does, underestimations can be expected.

Figure 40 illustrates the forecast of a whole week for tests 0 and 10. Due to the long time scale in the figure it is hard to accurately evaluate the peak performance, but it is seen that test 10 captures the behaviour of the flow quite closely, both regarding base flow and the value of the peaks. Test 0 appears to follow the general behaviour of the flow but struggles with the predictions of the peaks.

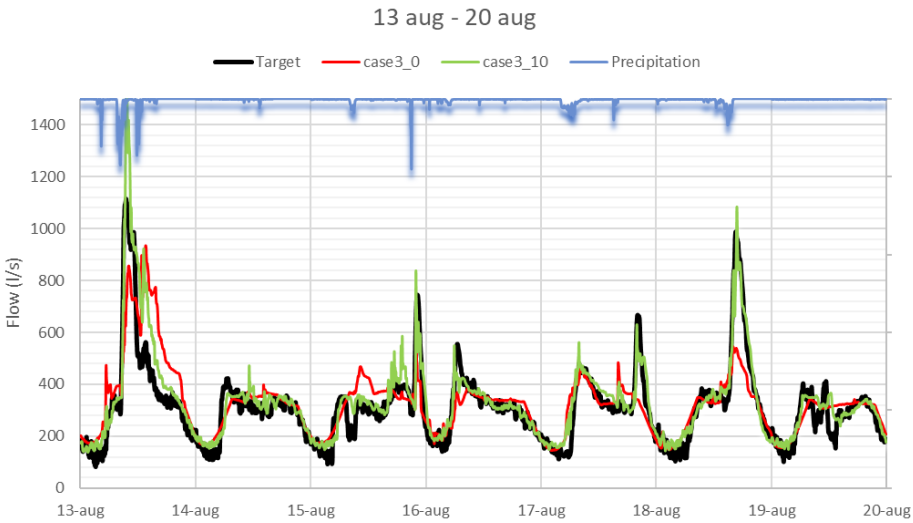

Figure 40: Comparison of the predictions from case 3.0 and 3.10 for 13th of August through 20th of August 2019.

4.2.7 Reference Case using Rain Gauges

Here, the performance of the tests using radar data is compared to the reference case which only uses rain gauge information, meaning that case 3 is compared to case 4. The cases are compared both with and without additional input variables. Test 9 is used instead of test 10, as test 10 includes rain gauge data.

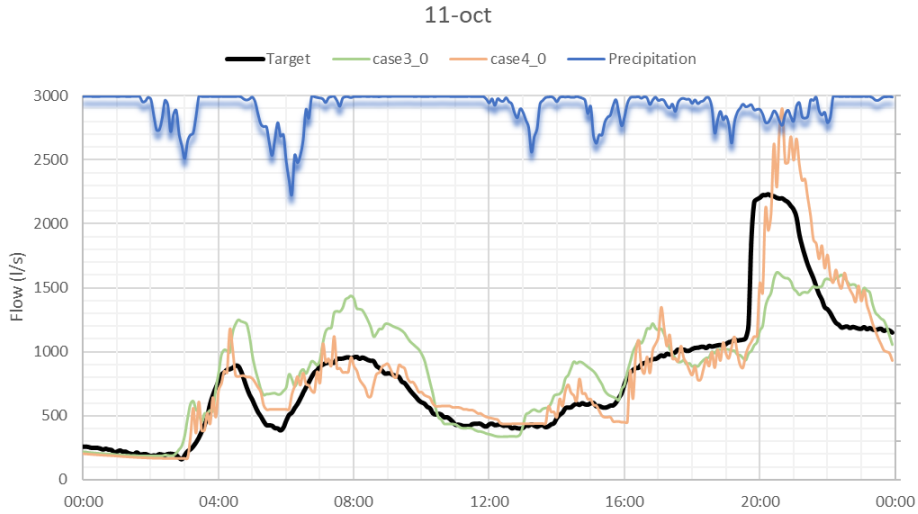

Figure 41: Comparison between case 3.0 (radar data) and 4.0 (rain gauge data) for the 27th of September 2019.

As seen in figure 41, the reference case with rain gauges performs a lot better than the radar case without additional input information added. The overall fit and value of peaks is more accurate than for the radar case. This could be due to various reasons but is likely related to issues with the radar data related to higher flows which makes the network perform poorly.

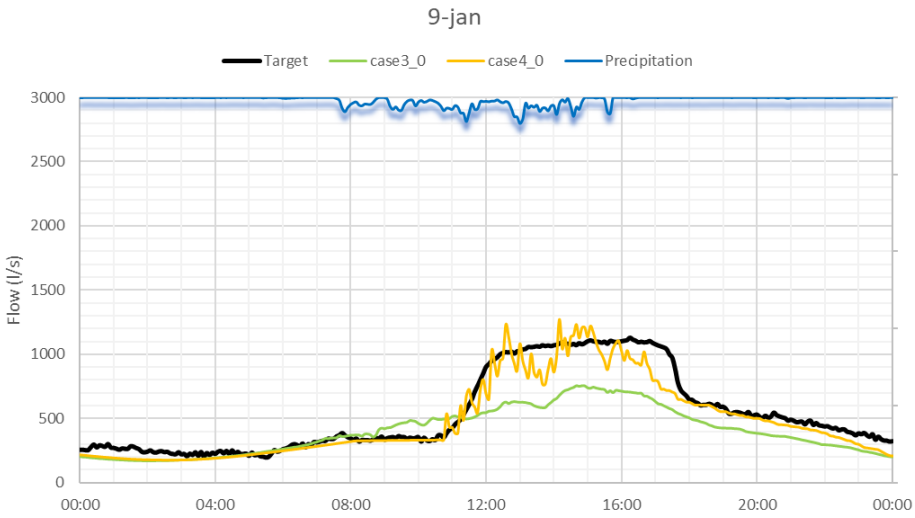

Figure 42: Comparison between case 3.3 (radar data, Dalby, Höje å), 3.9 (radar data, Dalby, Höje å, Källby WWTP) and case 4.9 (rain gauge, Dalby, Höje å, Källby WWTP) for the 11th of October 2019.

The 9th of January, seen in figure 42, the model using radar data fails to predict the increase in flow while the rain gauge based model fairly accurate describes the flow. The forecast from case 4.0 is volatile, oscillating up and down, indicating that it is very responsive to the precipitation information.

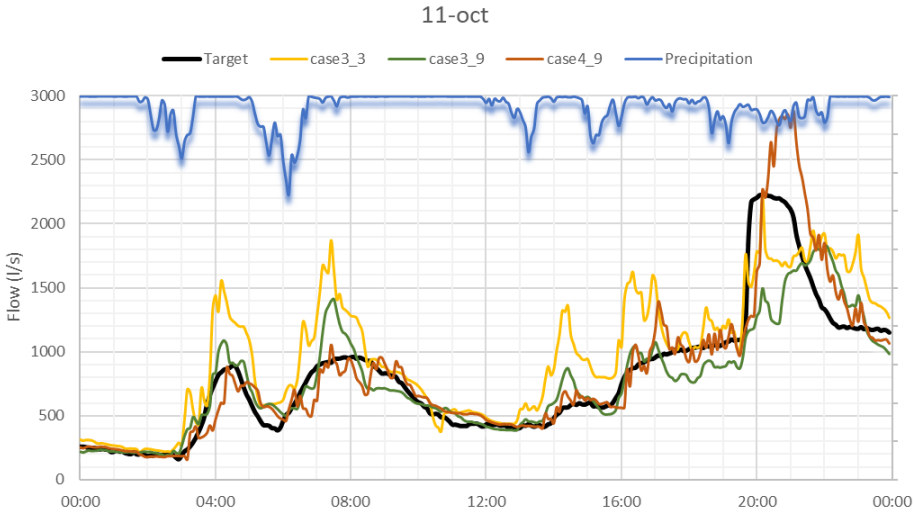

Figure 43: Comparison between case 3.3 (radar data, Dalby, Höje å), 3.9 (radar data, Dalby, Höje å, Källby WWTP) and case 4.9 (rain gauge, Dalby, Höje å, Källby WWTP) for the 11th of October 2019.

Adding additional information improves the performance of the cases using radar data. However, the difference between the different tests can sometimes be large, see figure 43. Test 3.3 does a better prediction of the peak at 20:00 while test 3.9 better captures the overall shape of the flow.

By adding information about the flow in Dalby, Höje å or Källby WWTP to the cases using radar data, their ability to forecast the flow improves greatly. This is not the case when using rain gauge information, as additional information only slightly improves the forecast. This might indicate that the rain gauge information alone is enough to make a decent forecast. The reason for this may be that the rain gauge data is consistent and reliable, even during large precipitation events, in contrary to the radar data. However, one must not forget the other benefits of the radar over the rain gauge, for example being able to measure precipitation with large spatial variability. Further, case 3.10, which earlier has been considered the best performing model, contains both rain gauge and radar data. This is an indication that the model further improves when it has access to both sources of precipitation information.

4.3 Optimisation of Network Structure

4.3.1 Hidden Layers

While investigating the optimal network structure for the purpose of forecasting, the first parameter tested is the number of nodes in the hidden layer. The reference number of nodes that is previously used in all tests are 8 nodes. The number of nodes tested are 1, 4, 16 and 100. Figure 44 illustrates the variation between runs of case 3 test 10 on 11th of October 2019. Case 3 test 10 includes both radar and rain gauge data and will further be used for all tests in this section.

The variations between the models are small and not systematic, making it difficult to conclude differences from the reference case, where the exception is the model with 1 node. The model with 1 node could not forecast a flow larger than 600 l/s, which indicates that one single node is insufficient for the model to yield a good forecast.

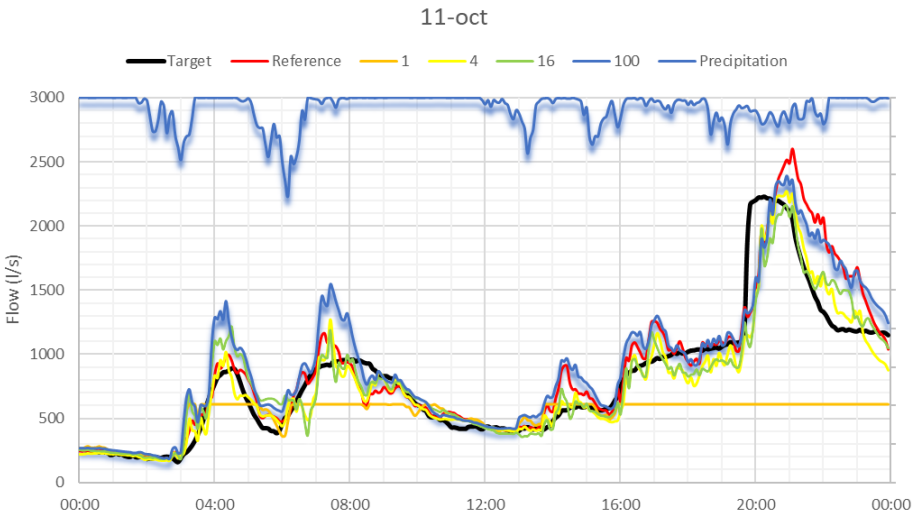

Figure 44: Comparison of different number of nodes in 1 hidden layer in Test 10. Reference curve has 8 nodes. The name of the curve is the number of nodes per hidden layer.

Secondly, the parameter studied is the number of hidden layers. It is concluded that increasing the number of hidden layers from one to two does not affect the performance of the model. Like for the tests with a single hidden layer, there does not seem to be any systematic improvement from the reference case when altering the number of nodes. Figure 45 shows the forecasts of models with two hidden layers containing 4, 8 and 100 nodes in each layer, respectively.

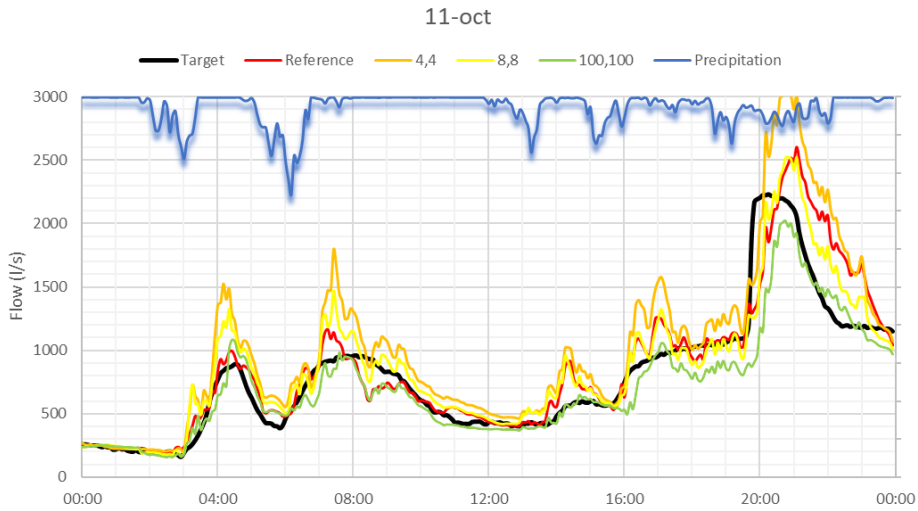

Figure 45: Comparison of different number of nodes and hidden layers in test 10. The reference has 8 nodes in 1 hidden layer. The name of the curve is the number of nodes per hidden layer.

It is concluded that for the purpose of forecasting wastewater flows based on several sources of input data, the neural network does not need to be more complex than a few hidden nodes in one layer. It is further concluded that added complexity will not, if at all, improve the model noticeably.

4.3.2 Roll Durations

When testing different roll durations, the reference case uses 15 min, 1 h, 3 h and 8 h, which is also what is used by all tests in section 4.1. Several different additional roll durations are tested, some longer and some shorter than the reference case. As the roll durations are supposed to simulate a short-term memory of the system, a comparison between case 3.0 (only radar) and case 3.10 (radar, rain gauge, Källby, Dalby, Höje å) is of interest to see how they react respectively.

Figure 46 and figure 47 show the performance of the model on October 11th for tests with different roll durations, for case 3.0 and 3.10 respectively. The performance of case 3.0 decreases drastically when no roll durations are used. This is likely because the model lacks information about recent rain events and the flow conditions in the sewer system, as well as about infiltration and runoff in the catchment. This can to some extent be replaced by the simulated memory from the roll durations. Case 3.10 is aided by the additional information from Dalby, Höje å and Källby WWTP, which makes up for the lack of roll durations. However, there is no significant difference

between any of the tests in figure 47, although the reference test might perform slightly better than the others.

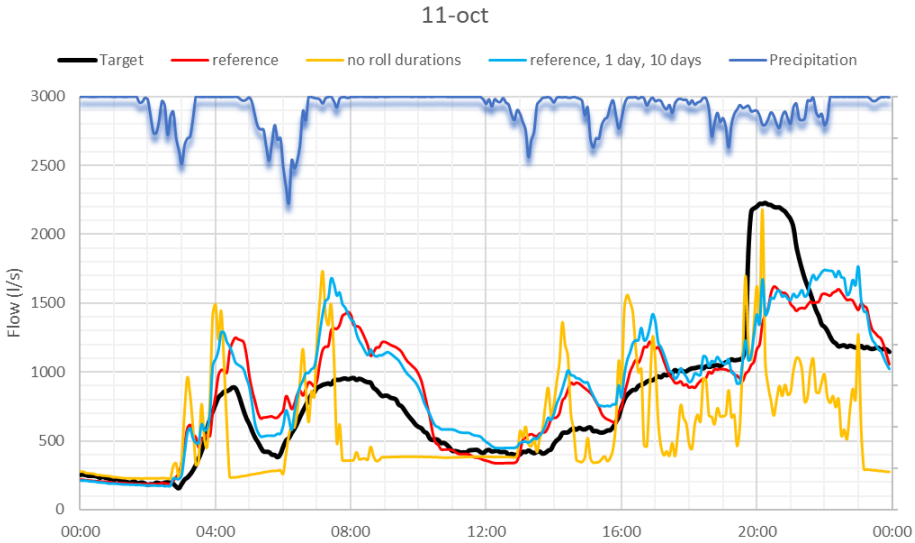

Figure 46: Test 0 with different roll durations.

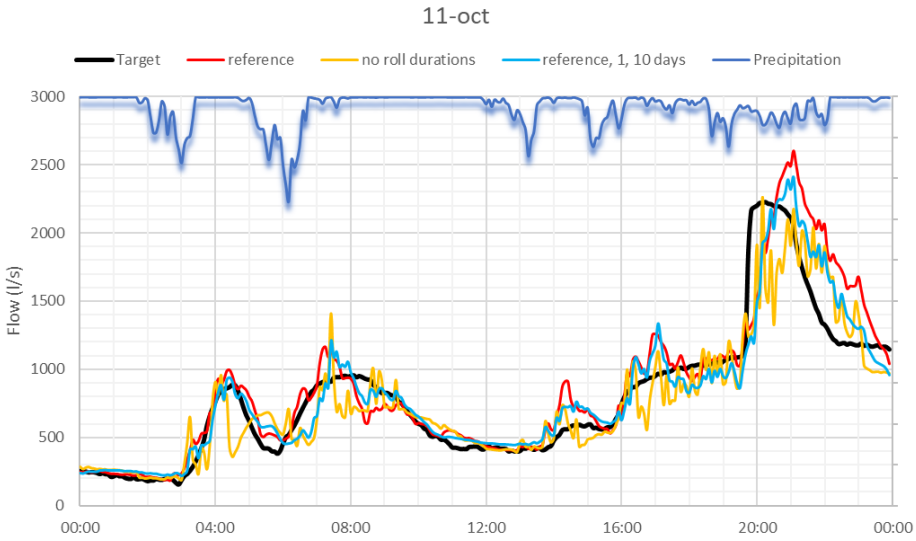

Figure 47: Test 10 with different roll durations.

4.4 Comparison of Error and Uncertainty

After evaluating the input variables and the structure of the network, it is desired to estimate the size of the error and uncertainty in the forecasts. It is interesting to evaluate this to know generally how surely the model can be trusted when put in use. The best performing tests, with and without rain gauge data, which are tests 9 and 10, are studied together with test 0 as a reference.

4.4.1 Systematic Error

By studying the error as a function of the flow, an estimation of the error for a certain flow can be made, illustrated in figure 48. The error is calculated as the forecasted value minus the corresponding measured value at Källby WWTP. The error seems to be systematic and follow a linear pattern, shown in Appendix C – Regression of systematic error, from which a linear regression is constructed yielding an average error as a function of the flow. The purpose of the dashed line, which represents the error for the time series cleared of measured flows smaller than 500 l/s, is to give a better picture of the error in the larger flow predictions, which are of main interest.

As seen in figure 48, test 10 has a significantly lower error than both test 9 and test 0, both with and without smaller flows. Test 9 shows a large difference between cleared flows and non-cleared flows, indicating that the error is larger for the larger flows and smaller for the lower flows. The error is negative for all flows larger than 750 l/s, meaning that the model is systematically underestimating all larger flows.

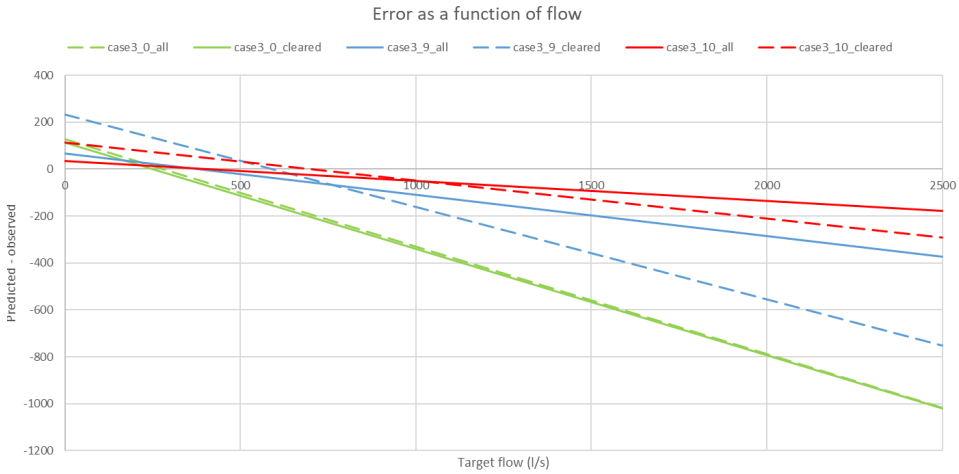

Figure 48: Comparison between the systematic error for the different tests, where the error is a function of the target flow. The dashed lines show the error for the same time series cleared of flows less than 500 l/s.

4.4.2 Standard Deviation

When observing the standard deviation in the forecasted flow, patterns can be discerned for how the standard deviation varies with increased flows. Figure 49 shows the systematic variation of the standard deviation for forecasted flows in three tests. In all three tests there is a clear linear increase in standard devation as the forecasted flow increases, but the spread and inclination decreases for models with more input variables. The linear relationship between forecast and standard deviation implies that the standard deviation for the forecast in a model is directly proportional to the value of the flow. For test 10, and to some degree test 9, there is a very sharp and clear inclination, especially visible for higher flows. As the inclination does not seem natural, it is suspected that it is related to how the neural networks model functions and produces the output.

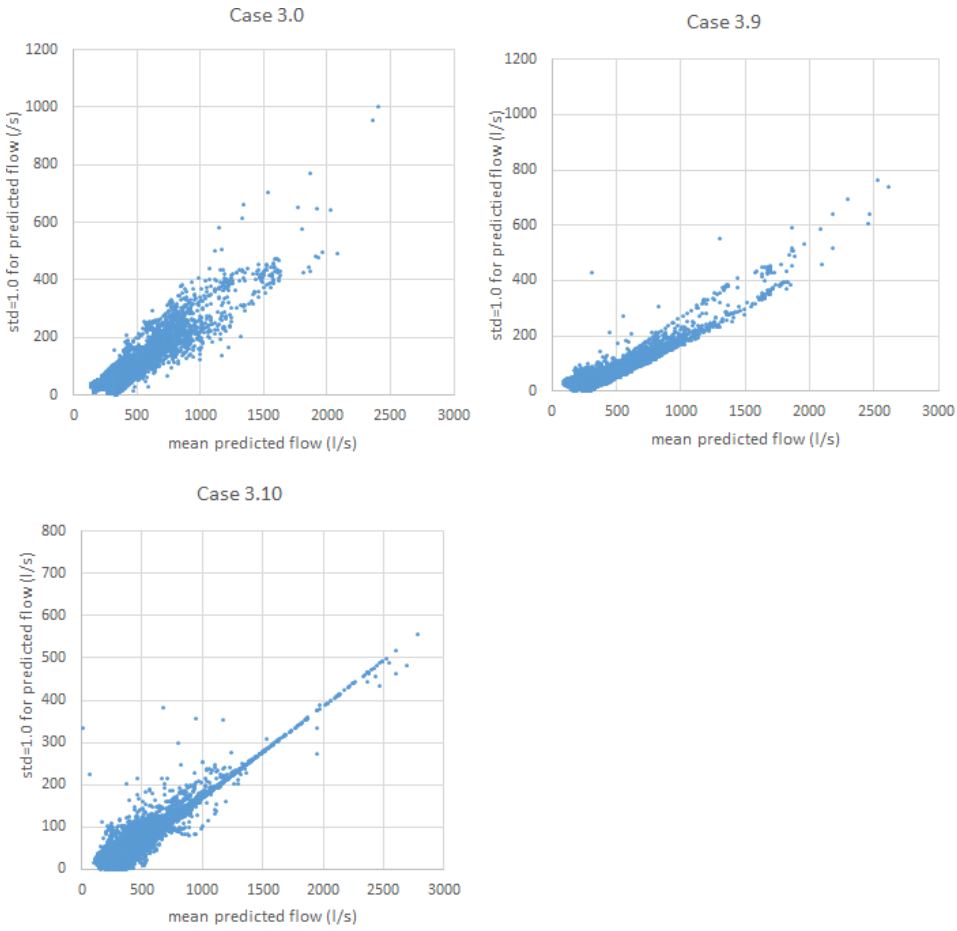

Figure 49: Plots showing how the standard deviation for forecasted flows increases.

4.5 Investigation of Prediction Time

4.5.1 Maximal Prediction Time

In this section, the maximal prediction time is investigated with the best performing model, which is case 3.10. It is concluded that as the prediction time increases, the forecasts gradually worsen and gets increasingly delayed, illustrated in figure 50. All prediction times tested show the same pattern in response to the precipitation, but with an increasing delay.

There is no major difference in the forecast of the 60-min prediction and the 45-min prediction. This indicates that using prediction times lower than 60 min does not drastically improve the performance of the model. This could further indicate that a 60-min forecast is somewhat reliable.

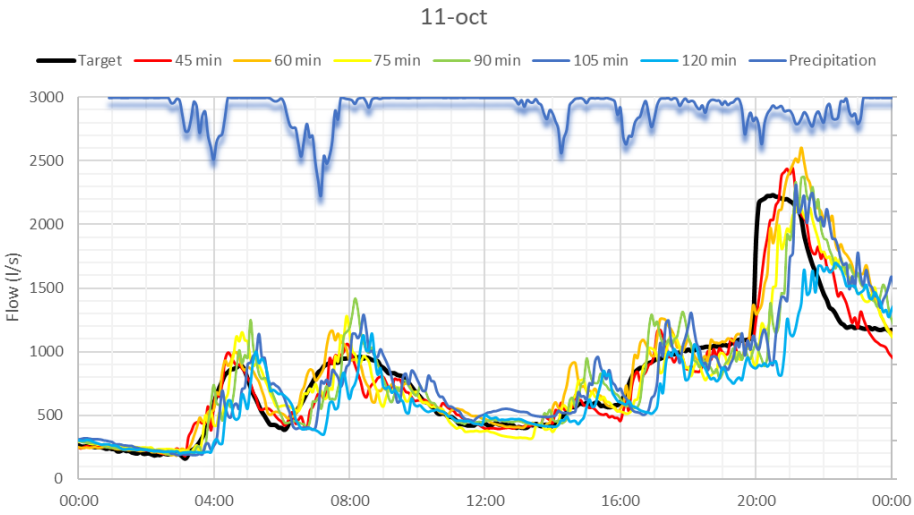

Figure 50: The different prediction times tested for the 11th of October 2019.

Figure 51 shows the precipitation event on the 6th of December 2019, where the difference between the different prediction times becomes even clearer. An increase from 60 to 75 min drastically increases the delay, indicating that 60 min might be the upper boundary of how far the prediction time can be made to yield reasonable results.

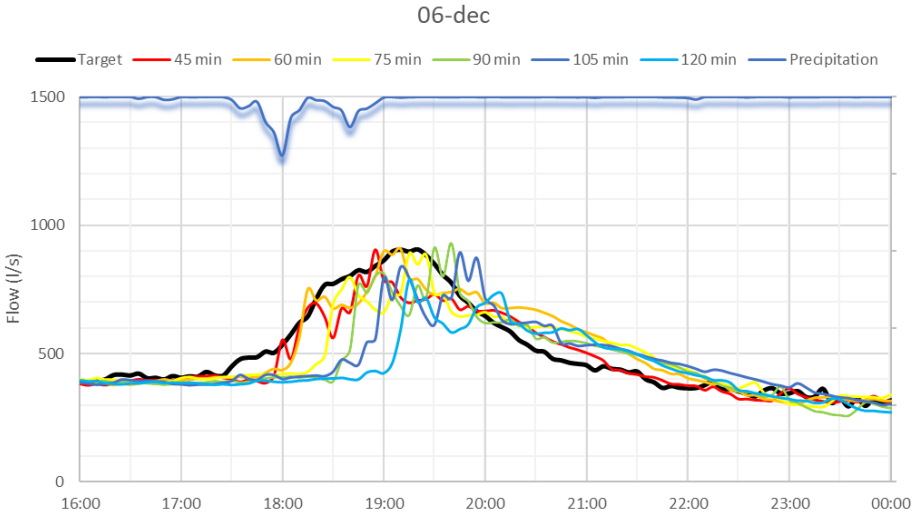

Figure 51: The different prediction times tested for the 6th of December 2019.

4.5.2 Further Extension of the Prediction Time

It is further investigated whether a perfect forecast of precipitation, here in form of rain gauge information put 4 h ahead of time, could be used to extend the forecast time of the flow to 4 h.

As seen in figure 52, the results from the 4 h prediction show that the network indeed is capable of making such a lengthy forecast, as both the timing and the value of the peaks are generally accurate. This means that if it is desired to extend the prediction time further than 1 h, some kind of forecast could be added as input to the neural network. Potentially, with a connected network of X-band radars or using a C-band radar, radar data from areas far away from Lund could be used as input data to produce this kind of forecast.

However, as the prediction is based on a perfect precipitation forecast, there are reasons to suspect that the neural network would perform worse in reality where the forecasts are uncertain. How this uncertainty affects the network is not known and could be further studied, but it is suspected to worsen the prediction. In conclusion, one could argue that the quality of the prediction depends on the reliability and quality of the forecast but that nonetheless, it is possible to make a forecast of the flow 4 h ahead. This could be further studied to increase the forecasting time potentially.

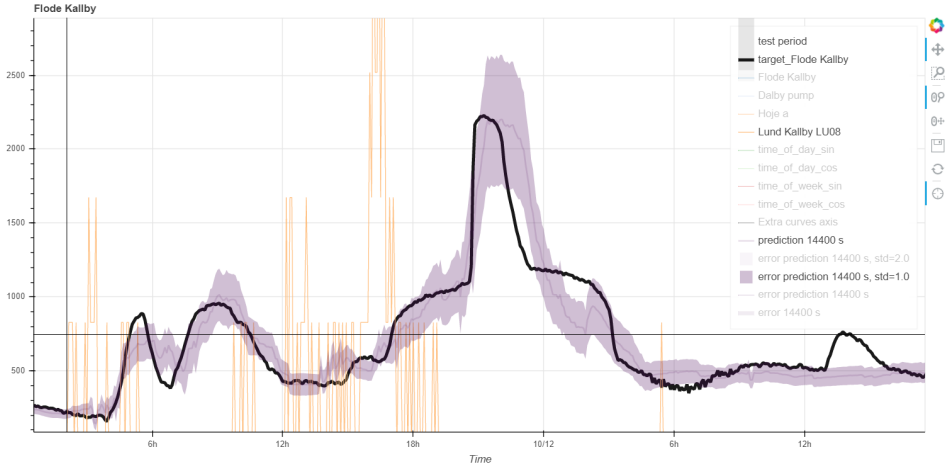

Figure 52: A 4h prediction using rain gauge information as a perfect forecast input for the 11th of October 2019. The black line is the measured flow at Källby WWTP and the thinner grey line is the prediction. The yellow lines illustrate precipitation and the purple field is the uncertainty in the prediction.

4.6 Comparison Neural Network and Conventional Model

It is worth stressing that a forecast made by the neural network is different from a simulation made in MIKE URBAN. The simulation uses rain events as input to retroactively simulate the reaction in the catchment and sewer system, whereas the neural network makes a 60 min forecast based on present conditions. The MIKE URBAN model can simulate flow time series for all different parts of Lund but only the simulated inflow into Källby WWTP is used. The purpose of this comparison is to see how the 60 min forecast compares with the simulation, in relation to the value measured at Källby WWTP.

While comparing the neural network with the conventional model, the best case and test is selected for the comparison which is case 3.10. It includes information from Höje å, Dalby, Källby, radar and rain gauges.

Illustrated in figure 53, the conventional model generally simulates the behaviour of the flow well, but tends to overestimate the flow of the peaks and gives an early response to precipitation. The forecast by case 3.10 closely follows the measured flow but is more volatile in its forecast compared to the smooth curve of the simulation from the conventional model.

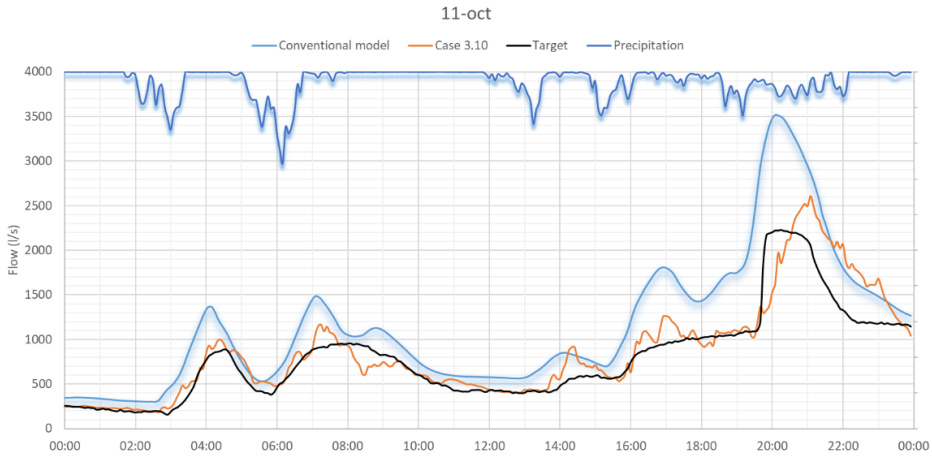

Figure 53: Simulated (conventional model) versus forecasted flow (case 3.10) for the 11th of October 2019.

Figure 54 gives another example of how the conventional model overestimates the flow. Case 3.10 overestimates the flow as well, but not to the same extent as the conventional model.

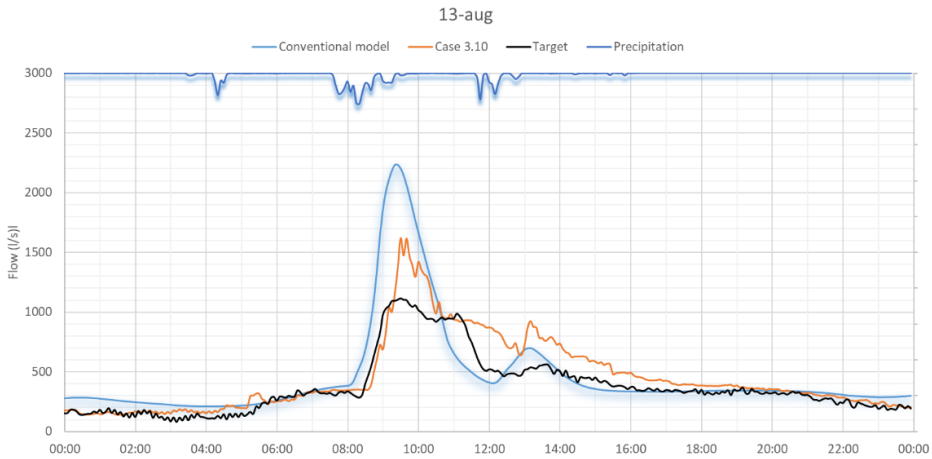

Figure 54: Simulated (conventional model) versus forecasted flow (case 3.10) for the 13th of August 2019.

Figure 55 illustrates that the conventional model sometimes struggles with simulating the base flow while case 3.10 gives a good forecast. Studying the peaks, both models make generally accurate reactions. However, the conventional model tends to overestimate the flow again.

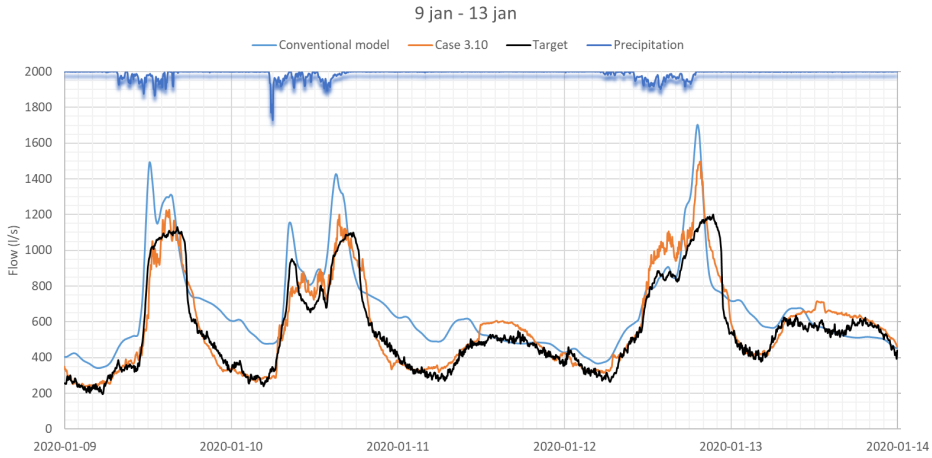

Figure 55: Simulated (conventional model) versus forecasted flow (case 3.10) from 9th of January to 13th of January 2019.

As seen in figure 56, the conventional model systematically overestimates the flow while case 3.10 systematically underestimates the flow. This becomes even clearer by studying the larger flows, as seen in figure 57. Figure 57 is cleared of flows smaller than 1 000 l/s, which shows how the two models

further deviates from the 1:1 line. The width of the cloud of forecasted and simulated values further show how varying the forecast or simulation can be for a certain target flow. The cloud is narrower for the neural network which indicates that the model is generally more accurate than the conventional model.

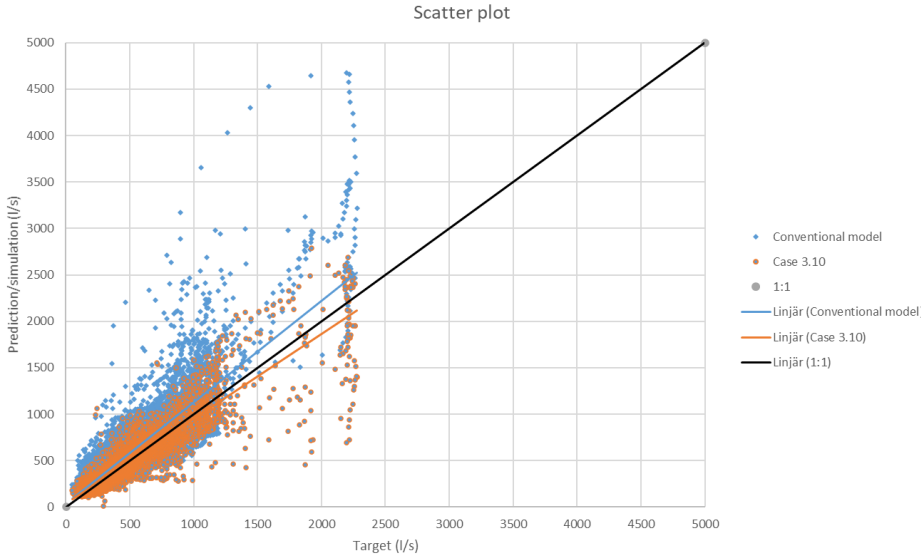

Figure 56: Scatter plot for the conventional model (blue) and case 3.10 (orange).

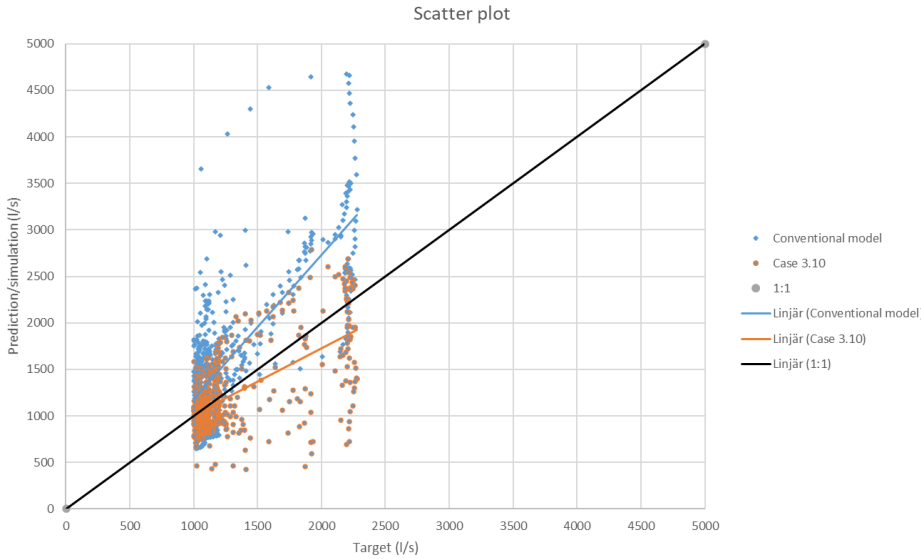

Figure 57: Scatter plot cleared of flows smaller than 1000 l/s for the conventional model (blue) and case 3.10 (orange).

5 Discussion

The finding that a neural network can be trained to make a hydrological forecast of a waste- and stormwater flow based on non-bias-corrected radar data is considered a success for this project. Requiring only a target, here in form of historical data of the combined wastewater flow, to be forecasted and precipitation information from a radar or rain gauge, a model like this could be set up for municipalities or communities with relatively little work. By adding additional information about the wastewater flow or hydrological conditions in the catchment, the accuracy of a forecast made by the model further improves greatly. The versatility of a neural network further gives it much potential as either an urban planning or flood warning tool.

However, if the model is to reach its full potential, the performance is dependent on the quality of the radar data. This is also the largest identified issue in the project, where perhaps the most important factor is the observed evidence of the attenuation effect during heavy rains which leads to no or reduced amounts of precipitation being recorded. This is unfortunate, as heavy rains are the most important events to study and train for the purpose of the project. Additionally, large amounts of radar data are missing because of various technical issues. This affects the model in the same way as effectively having a shorter time period for training the neural network.

Today, there is no model forecasting the incoming flow at Källby WWTP in real time. To be able to forecast the flow 1 h ahead of time, with just a few input variables and a neural network would be valuable information for the WWTP. For future studies, it would be highly interesting to investigate which variables could be added to the network to extend the prediction time further to 4 – 6 h. This could be done by testing the potential of a network of radars or include information from rain gauges and wind in areas far away from Lund, together with the neural network.

Comparing the cases using radar data with the reference case using only rain gauge information, it is seen that the best radar case, which also uses rain gauge data, slightly outperforms the reference case. This means that depending on use, purpose and desired accuracy of the forecast, using only rain gauge information may be good enough. However, accuracy of the radar would likely be better if the radar-related issues, which cripple the accuracy of the radar, could be solved. The issues could be reduced in a variety of ways, for example by using a network of radars, by performing bias-correction, or by using different levels of radar data. However, radar and rain gauges may well be used together as they both have their strengths and weaknesses and can further complement each other.

The radar data used in the project is not bias-corrected, as the neural network itself performs a kind of bias-correction when it adjusts the weights connected to different inputs. This is a great benefit of a neural network approach since a bias-correction, which requires extensive work, is usually necessary for a program to yield satisfying results. However, the neural network could possibly perform even better if a rough correction of the data already is done before it enters the network.

Due to how the neural network is trained on historical catchment data, continuous changes in the wastewater system or catchment would decrease the accuracy of the model, as the relationship between input and output would change. Significant and sudden changes in the system could potentially make a neural network useless and require a new period of training with new data. It is therefore recommended to keep the model up to date by training it on as new data as possible.

Finally, this study should be treated as a first indication of the potential to forecast waste- and stormwater flow with a neural network. Since a somewhat subjective method of analysis is used, a more sophisticated study could be carried out to better quantify the discoveries and further cement the conclusions. It is also advised to use a testing period to further evaluate on an additional independent data set.

Sources of Error

The evaluation of the different cases, tests and comparisons were not done on an independent set of test data. The lack of a test data set could be viewed as problematic, as it is not known how the model performs on data that is not included in training and validation. However, concerning the main aim of the study, which is to determine the influence and importance of other physical variables together with the radar data, it is considered that comparisons still can be made between the different models.

The same line of reasoning could be applied to the choice of validation period, which for this study is set to December and January. These months are considered representative and similar enough to any other months regarding precipitation and runoff for an urban environment like Lund. A validation on different seasons could be beneficial, but as the study is focused on an urban area with no considerable snow melt or other seasonal precipitation patterns, the selection of training and validation period is considered to have a minor impact on the result.

Ideally, the measured flow at Källby WWTP is a function of only the waste-

and stormwater that naturally flows into the plant. However, this is not the case, as there are technical complications and measures taken by operators that affect the inflow. For example, there are several pumping stations around the sewage system that can regulate the flow. Information about how and when these stations operate were not available for this project. Additionally, Källby WWTP cannot measure any flows larger than 2 200 l/s and there are a few events in the training process where this occurs. Lastly, as seen in figure 39, there is a pattern forming at 1 200 l/s where the flow is not exceeding this value. The reason for this soft cap is not known but could possibly be due to flow restrictions at the treatment plant made to prevent the flow from exceeding a certain limit. These issues will affect the training of the network as it is very difficult, if not impossible, for the network to consider these unnatural flow variations in the training process.

Since radar data for large areas of Lund is missing because of the noise filter, it would be interesting to evaluate the performance of the network if noise free radar data for all of Lund could be included. The reason for the noise is unknown, but it could be related to tall buildings or pollution particles in the air over the eastern parts of Lund. Bad data issues are important to identify and address before the data is used for training. A sign of this is the relatively poor performance of case 1, where identified problem areas were not removed.

Representing many aggregated grid points of radar data with one value for the entire area is convenient but could be problematic. If some of the points are malfunctioning, by for example vastly overestimating the precipitation, the network does not know which these points are as it only treats the aggregated value for the entire area. This could have a large impact on the value for the polygon, especially if the points malfunction in inconsistent manners. If the network instead could use each grid point as separate input variables, it could adjust the weight connected to each point. This would give each grid point an input node which could be weighted individually in the neural network, and not just one input node for the entire polygon.

Before evaluating the usefulness of other data sources for the neural network, a brief inspection of the data was made after which it was considered that the quality of the data is decent. Nevertheless, there is still a possibility that there are inaccuracies in the data. As a neural network is supposed to be capable of determining what data is useful and not, issues related to data quality should not be of large importance. However, as removing malfunctioning areas of radar data improved the results, further treatment seems to be beneficial for the model.

The time period which the network is trained on stretches from May 2019 – Jan 2020, which is a relatively short period of time and a limitation to the project. The neural network benefits from large data series to use for the training process and produce an optimal model. As the time period does not encompass several years, it cannot capture eventual seasonal patterns in both the wastewater flow and the precipitation. Therefore, if available, a longer time series of data would likely have improved the model.

In this study, it is assumed that each value for every input variable has been designated a time stamp in a consistent and synchronised manner but this might necessarily not be the case. As the clocks determining the time for each input variable may drift over time, the relationship in timing between variables may change. The risk of desynchronisation will increase with an increasing amount of data variables being used as well as for longer time periods. For example, two clocks that initially were synchronised might after 3 months measure the time with a difference of 5 minutes. If the incorrect clock then is adjusted to the correct time, it will have an impact on the data time series. This change could confuse the network which could degrade the result.

Future Studies

This study is limited to only looking at variables connected directly to Lund, which may be why forecasts further than 60 min ahead could not be made satisfactory. To achieve a model with longer prediction times, radar data from areas outside of Lund would likely be required. In such a case, the wind speed and wind direction together with rain gauge information from these areas could also be useful information for the network.

The comparison between case 2 and case 3 showed that dividing the radar data into sub-catchments made negligible improvements to the performance of the model. This could be related to the 500 by 500 m resolution restriction of the radar but it is more likely that for this kind of application and purpose, which is to forecast combined sewer flow, the aggregated precipitation for the entirety of Lund provides good enough information to the network. However, the removal of noisy data in eastern Lund made a large positive impact on the performance of the model.

Opposite to increasing the number of sub-catchments, a less detailed division of catchment could be tested. For example, aggregation of radar data in a large circle or square around Lund could be investigated. If the radar data aggregated in such a crude manner proves to be equally or even more useful for the neural network, it would further indicate that no consideration to

the hydrological boundaries needs to be taken for this kind of models.

What was not included in this study, but could be of interest for future studies, is to develop a model that calculates the probability of exceeding a certain forecasted flow at a given time. A model like this could be programmed to give a warning whenever a certain flow is believed to be reached soon with a certain probability. This could be used as a safety measure which always would be running in the background at for example, Källby WWTP.

6 Conclusion

The purpose of the study was to evaluate the use of X-band radar data with a neural network. The study concludes that it is possible to train a neural network with non-bias-corrected radar data over Lund to forecast the flow at Källby WWTP up to 60 min ahead. Prediction times longer than 60 min degrade the result and cause a delay in the forecast. It was further indicated that the prediction time could potentially be increased by adding forecasted or measured precipitation information from areas outside of Lund.

A model using only rain gauge information performed better than a model using only radar data, which is believed to be mainly related to attenuation issues in the radar data. As a rain gauge provides more consistent data for high-intensity rainfalls, it might be more reliable for the neural network. However, a model combining both radar and rain gauge data performs even better as it benefits from both the spatially high-resolution radar and the consistent performance of the rain gauge, especially when attenuation is affecting the radar.

The model generally forecasted base and medium flows of up to 1 000 l/s well but struggled when the flows exceeded 2 000 l/s, which is probably mainly related to attenuation issues connected to larger precipitation events. Further, the model consistently underestimated the target flow slightly and the error in the forecast was proportional to the value of the forecasted flow, resulting in large error and uncertainty for the largest flows.

For the purpose of forecasting, aggregating the radar data in sub-catchments instead of one large catchment made little, if any, improvements to the result. However, removing areas with problems connected to the radar data was important for the performance of the model.

The complexity of the network structure showed to be of no importance for the quality of the forecast, indicating that a simple network with one hidden layer with two or more nodes is sufficient to produce an accurate forecast. The performance of the model was improved by adding more input variables, and the combination of variables that gave the best performance was:

- X-band radar data
- flow information from sewer system outside of radar coverage (Dalby)
- discharge in receiving watercourse for the catchment (Höje å)
- rain gauge information
- current flow at the WWTP (Källby)

Groundwater and wind information did not improve the forecasts. The changes in groundwater level are assumed to be a process too slow to be of value for a model of this purpose. The wind information could potentially be beneficial if it is complemented with data from nearby rain gauges.

It was possible to make a forecast by using only radar data, although it yielded significantly worse results without the other input variables. However, depending on the purpose and demands of the user, a forecast like this might still be useful. If no input variables apart from radar data are available, it is possible to improve the forecast by adding rolling averages on the input, which functions as a short-term memory of the catchment within the model.

Another objective of the study was to compare the model with a conventional physics-based model. It was concluded that the predicted values of a 60 min forecast made by the model is comparable to a simulation by a conventional model, indicating that the forecast is not worse than the simulations from a MIKE URBAN model being in use today. However, one needs to remember that the models have different purposes and this comparison should only be considered with much care.

7 References

Agarp, A. F. (2018). 'Deep Learning using Rectified Linear Units (ReLU)', arXiv:1803.08375

Alemu, H. Z., Wu, W., Zhao, J. (2018). 'Feedforward Neural Networks with a Hidden Layer Regularization Method', Symmetry, pp. 525 – 543. doi:10.3390/sym10100525

Bashide, M. (2015). *Modelling of Käppala Wastewater Treatment Plant.* Master thesis. Lund University. http://lup.lub.lu.se/luur/download?func=downloadFilerecordOId=8046524fileOId=8046525 (Accessed: 23 January, 2020)

City of Salem. (2020). *Inflow and Infiltration.* https://salemva.gov/Departments/Water-Sewer-Dept/Wastewater-Operations/Inflow-and-Infiltration (Accessed: 24 February, 2020)

DHI. (2017). *Mike Urban - Collection System. Modelling of storm water drainage networks and sewer collection systems.* https://manuals.mikepoweredbydhi.help/2017/Cities/CollectionSystem.pdf (Accessed: 14 May, 2020)

Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N-E., Quirmbach, M., Vaes, G., Vieux, Baxter. (2004). 'Towards a roadmap for use of rada rainfall data in urban drainage', Journal of Hydrology, pp. 186 – 202. doi: 10.1016/j.jhydrol.2004.08.004

Ellis, J.B. (2001). 'Sewer Infiltration/exfiltration and Interactions with Sewer Flows and Groundwater Quality. Interactions between Sewers, Treatment Plants and Receiving Waters in Urban Areas – INTERUBA II', pp. 311 – 319, Lisbon

Furuno. (2020). *WR-2100 Brochure.* https://www.furuno.no/Userfiles/Sites/files/WR-2100_E.pdf (Accessed: 23 January, 2020)

Gurney, K. (1997). *An introduction to neural networks.* https://www.inf.ed.ac.uk/teaching/courses/nlu/assets/reading/Gurney_et_al.pdf (Accessed: 28 January, 2020)

Gustafsson, L. G. (2000). 'Alternative Drainage Schemes for Reduction of Inflow/Infiltration - Prediction and Follow-Up of Effects with the Aid of

an Integrated Sewer/Aquifer Model', pp. 21 – 37

Hardesty, L. (2017). *Explained: Neural networks.* http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414 (Accessed: 29 April, 2020)

Hill, T., Marquez, L., O'Connor, M., Remus, W. (1993). 'Artificial neural networks for forecasting and decision making', International Journal of Forecasting, pp. 10(1): 5 – 15. doi: 10.1016/0169-2070(94)90045-0

Hedell, C., Kalm, A. (2019). *Applying X-band radar data in urban hydrology - Adjusting data for a neural network model, based on the pilot project in Dalby 2018.* Lund University.

Johansson, B., Chen, D. (2003). 'The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modelling', International Journal of Climatology, pp. 1523 – 1535. doi:10.1002/joc.951

Karpf, C., Krebs, P. (2011). 'Quantification of groundwater infiltration and surface waterinflows in urban sewer networks based on a multiple model approach', Water Research, pp. 3129-3136. doi: 10.1016/j.watres.2011.03.022

Lengfeld, K. (2016). 'A Simple Method for Attenuation Correction in Local X-band Radar Measurements Using C-band Radar Data', Journal of Atmospheric and Oceanic Technology, pp. 2315 – 2329. doi: 10.1175/JTECH-D-15-0091.1

Lengfeld, K., Clemens, M., Münster, H. and Ament, F. (2014). 'Performance of high-resolution X-band weather radar networks – the PATTERN example', Atmospheric Measurement Techniques, pp. 4151 – 4166. doi:10.5194/amt-7-4151-2014.

Lidström, V. (2020). *Vårt Vatten.* Solna: Svenskt Vatten AB

Lunds Kommun. (2018). *Lunds Kommuns översiktsplan del 1: Planstrategi.* https://www.lund.se/globalassets/lund.se/traf_infra/oversiktsplan/lunds-kommuns-oversiktsplan-del-1-planstrategi-antagen.pdf (Accessed: 29 April, 2020)

Naturvårdsverket. (2010). *Rening av avloppsvatten i Sverige.* http://www.naturvardsverket.se/Documents/publikationer6400/978-91-620-8629-9.pdf

(Accessed: 4 February, 2020)

Olsson, L. (2019). *Flow simulation in MIKE URBAN based on high-resolution X-band radar data - A case study in Lund.* Lund University.

SMHI. (2017). *Radar.* https://www.smhi.se/kunskapsbanken/meteorologi/radar-1.5942 (Accessed 14 May, 2020)

Solomatine, D., See, L.M., Abrahart, R.J. (2008). 'Data-Driven Modelling: Concepts, Approach and Experiences', Practical Hydroinformatics, pp. 17 – 30. doi:https://doi.org/10.1007/978-3-540-79881-1_2

South, N., Hashemi, H., Olsson, L., Hosseini, S. H., Aspegren, H., Larsson, L., Berndtsson, R., Das, R., Marmbrandt, A., Olsson, J., Persson, A. (2019). *Väderradarteknik inom VA-området – test av metodik.* https://www.svensktvatten.se/contentassets/37084ad45b77468d9162f2044e8de473/sv-raprt-19-3.pdf (Accessed: 18 February, 2020)

Strangeways, I. (2010). 'A history of rain gauges', Royal Meteorological Society, pp. 133 – 138. doi: 10.1002/wea.548

Svenskt Vatten. (2016). *Avledning av dag-, drän- och spillvatten.* http://vav.griffel.net/filer/P110_del1_web_low

VA Fälttjänst AB (2018). *Hydrodynamisk analysrapport för 12 st mätpunkter för mätkampanj utförd i Lund sommaren/hösten 2018*

VA Syd (2018a). *Dagvattenplan för Lunds kommun.* https://www.vasyd.se/-/media/Documents/Informationsmaterial/Vatten-och-avlopp/Dagvatten/Dagvattenplan-fr-Lunds-komun-2018.pdf (Accessed: 22 January, 2020)

VA Syd (2018b). *Väderradar ska förutspå regn.* https://www.vasyd.se/Artiklar/Nyheter/Dagvatten/Vaderradar-20180705 (Accessed: 22 January, 2020)

VA Syd (2018c). *Åtgärdsplan för Lunds avlopp 2018.* https://www.vasyd.se/-/media/Documents/Informationsmaterial/Vatten-och-avlopp/Atgardsplaner-for-avlopp/tgrdsplan-Lund-2018-ver-2.pdf (Accessed: 22 January, 2020)

VA Syd (2019a). *Källby avloppsreningsverk.*

https://www.vasyd.se/Artiklar/Avlopp/Avloppsreningsverk/Kallby-avloppsreningsverk
(Accessed: 22 January, 2020)

VA Syd (2019b). *Sjölunda är navet i en regional lösning.*
https://www.vasyd.se/Artiklar/Avlopp/Hallbar-avloppsrening/Sjolunda-ar-navet
(Accessed: 22 January, 2020)

Wu, Y., Liu, L., Bae, J., Chow, K-H., Iyengar, A., Pu, C., Wei, W., Yu, L., Zhang, Q. (2019). 'Demystifying Learning Rate Policies for High Accuracy Training of Deep Neural Networks'. arXiv:1908.06477

Ying, Xue. (2018). 'An Overview of Overfitting and its Solutions', Journal of Physics, pp. 1 – 6. doi:10.1088/1742-6596/1168/2/022022

Zhang, G. P., Patuwo, E., Hu, M. Y. (1998). 'Forecasting with artificial neural networks: the state of the art', International journal of forecasting, pp. 14(1): 35 – 62. doi: 10.1016/S0169-2070(97)00044-7

Zhang, G. P. (2012). 'Neural Networks for Time-Series Forecasting', Handbook of Natural Computing, pp. 461-477. doi:10.1007/978-3-540-92910-9_14

Zhang, S., Song, Z. (2019). 'An ethic costumes classification model with optimized learning rate', Proceedings of SPIE, doi.org/10.1117/12.2539608

8 Appendix

Appendix A - Visualisation of variation between identically trained neural networks

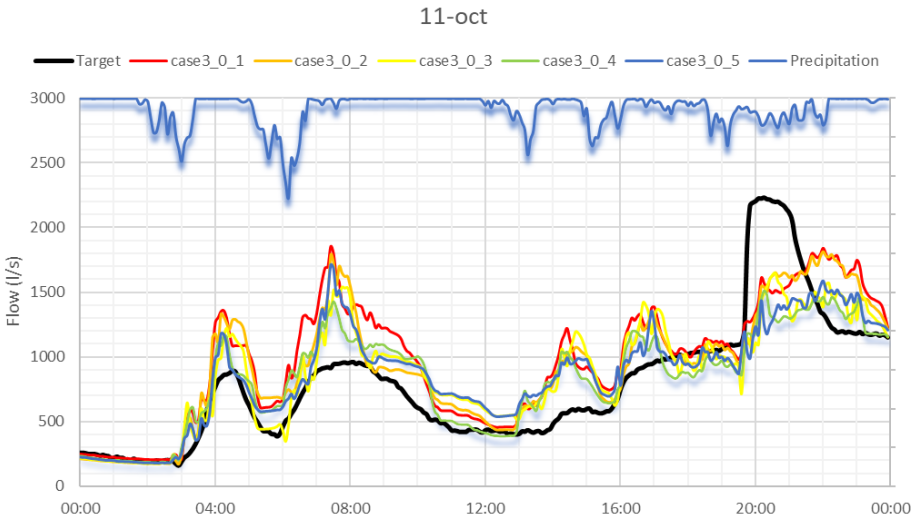

Figure 58: 5 runs of same case and test.

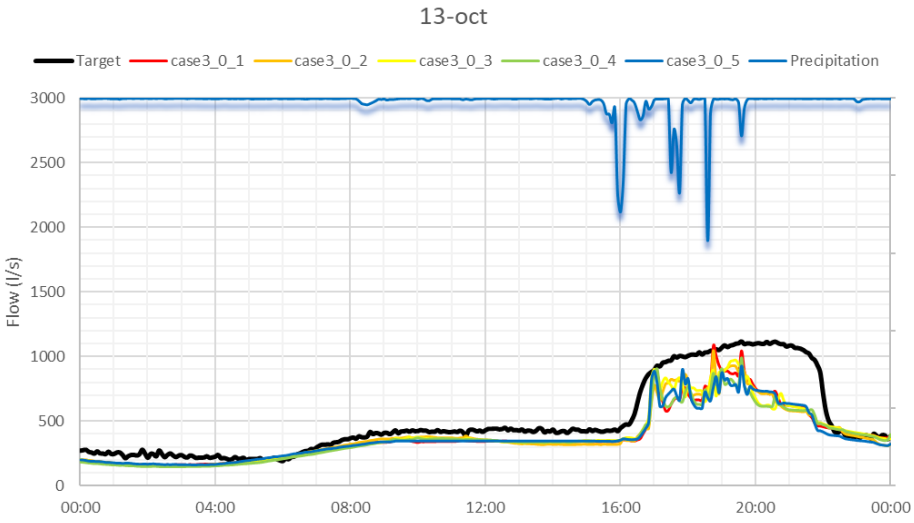

Figure 59: 5 runs of same case and test.

Appendix B - Accumulation of precipitation in grid points

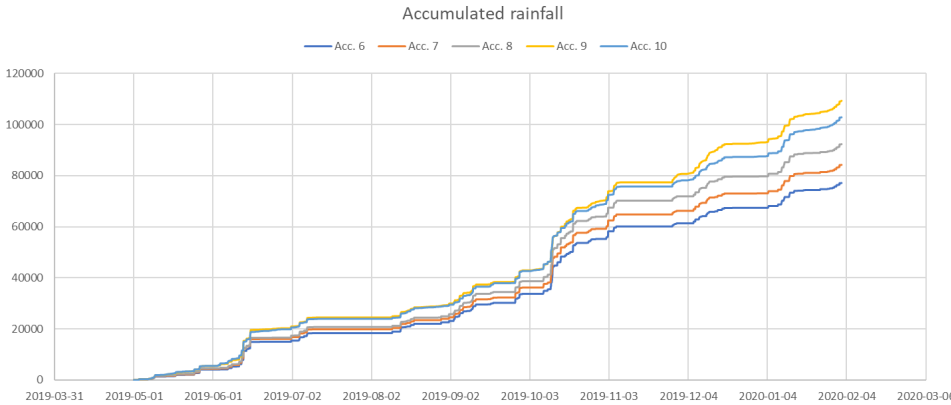

Figure 60: Accumulated rainfall measured by the radar for points 6 – 10.

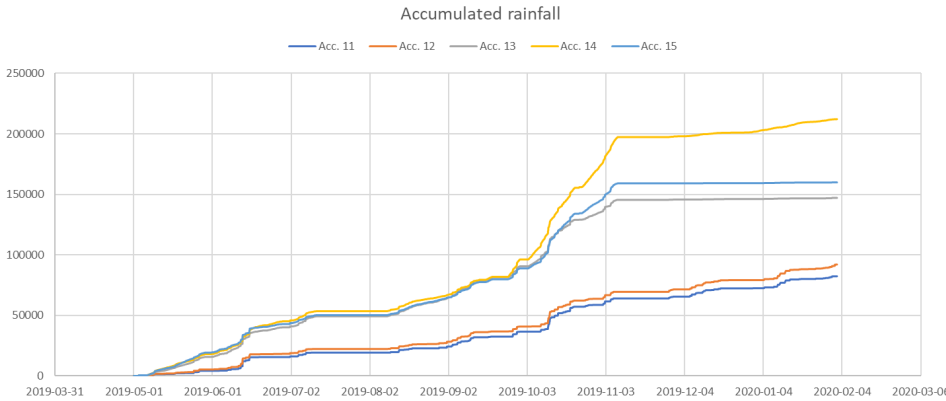

Figure 61: Accumulated rainfall measured by the radar for points 11 – 15.

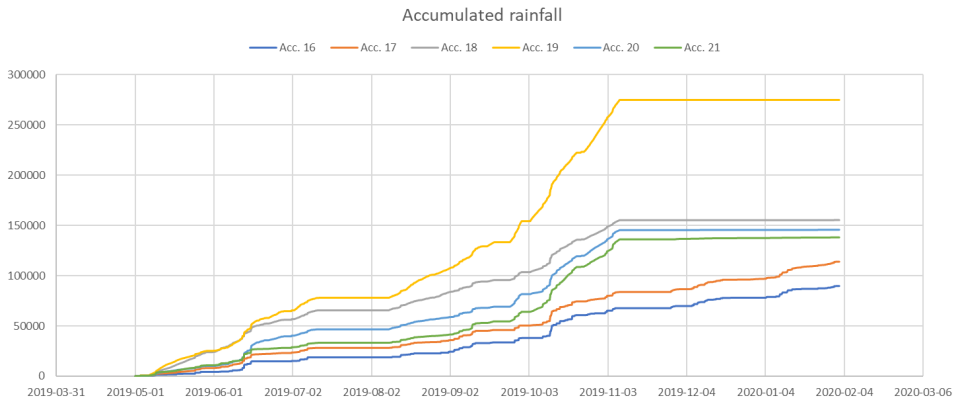

Figure 62: Accumulated rainfall measured by the radar for points 16 – 21.

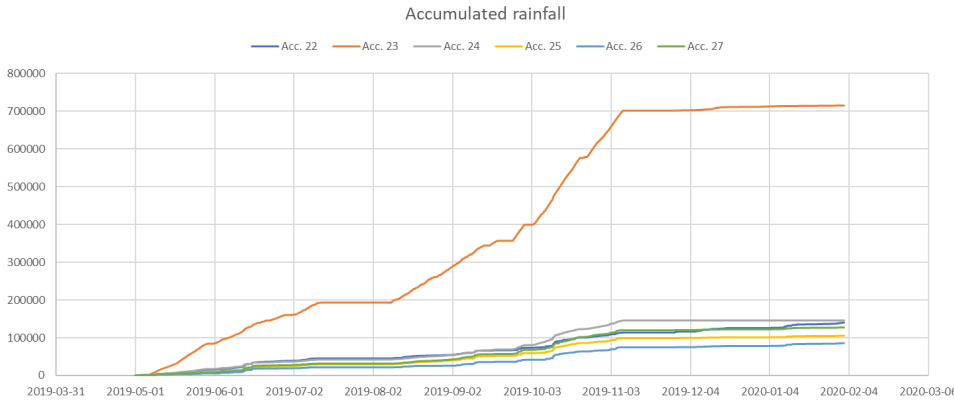

Figure 63: Accumulated rainfall measured by the radar for points 22 – 27.

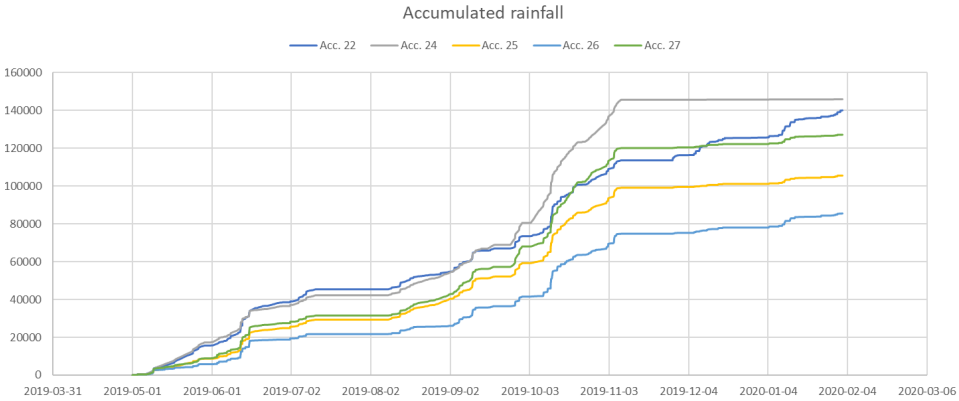

Figure 64: Accumulated rainfall measured by the radar for points 22, 24, 25, 26 and 27.

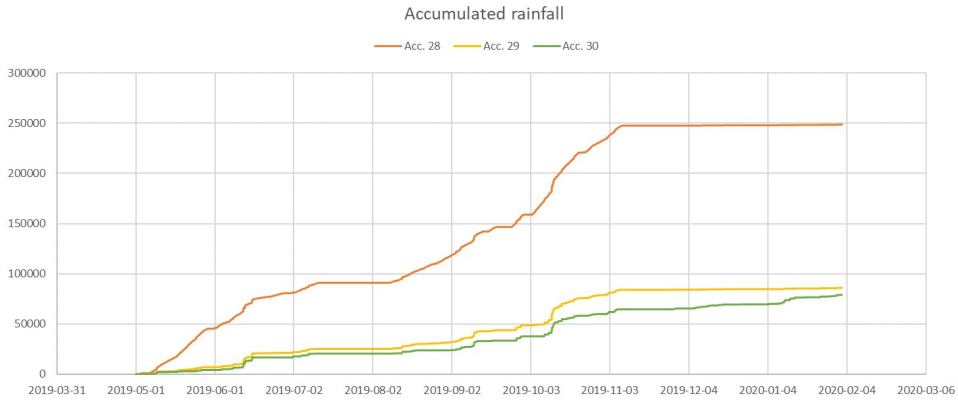

Figure 65: Accumulated rainfall measured by the radar for points 28 – 30.

Appendix C - Regression of systematic error

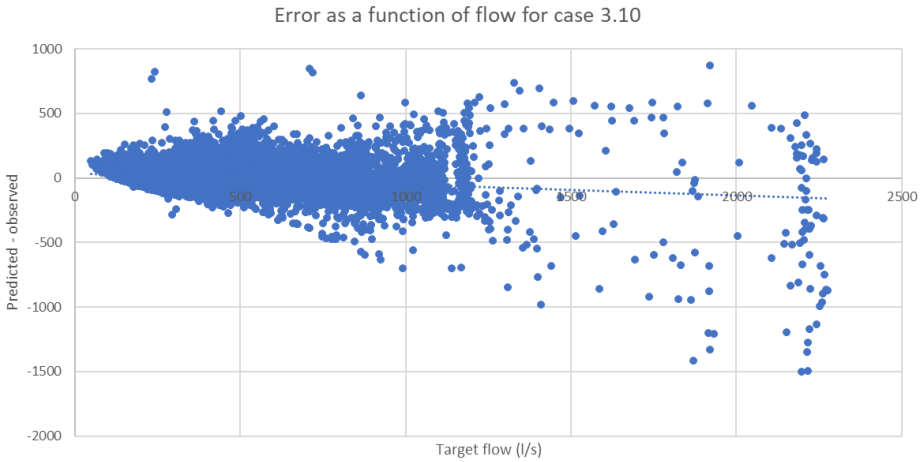

Figure 66: The systematic error plotted against the target flow for case 3.10.