University of Twente

Exploring the use of surrogate models to reconstruct historic discharges

Artificial neural networks for the reconstruction of the 1809 flood event of the Rhine river delta

Yorick Fredrix
26-Sep-18

Exploring the use of surrogate models to reconstruct historic discharges

Artificial neural networks for the reconstruction of the 1809 flood event of the Rhine river delta

Author

Yorick Fredrix

Under guidance of:

Ir. A. Bomers
Daily supervisor
University of Twente, Water Engineering Management

Dr. R.M.J. Schielen
Committee member
University of Twente, Water Engineering Management / Rijkswaterstaat

Prof. Dr. S.J.M.H. Hulscher
Head of graduation committee
University of Twente, Water Engineering Management

Abstract

Currently over the last 120 years discharges are measured in the Dutch river delta. These values are used to consider the possible discharge of once every 1.000 -10.000 year. This requires extrapolation of these 120 data points. By reconstructing historic floods, additional extreme values can be added to this dataset. Using these cases will reduce the uncertainty in what will occur once every 1.000 - 10.000 year.

Such a reconstruction can be made with the help of physical models. A physical model of the geophysical situation around the event can be made. In these models however multiple parameters remain unknown, most noticeable the discharge and roughness values. Whilst having multiple unknown parameters, standard calibration methods fail. Due to the complex calculation nature of these models, the calculation time constrains the use of multivariable optimization. In this regard meta models might have a solution. A meta model is a simpler model that represents the detailed hydrodynamic model. Within meta models there are two options, namely lower fidelity modelling and data modelling. The lower fidelity modelling is still a physical model with less details, like a coarser grid. Data modelling leaves all physical relations behind and tries to find relations between the input and output of the original model. In this thesis only data modelling is considered as it has the potentially largest speed increase.

For the reconstruction of the flooding of 1809, a detailed 2D model with the use of the software D-Flow FM is built. This model describes the geophysical parameters accurately and has a range of parameters for the unknown discharges and roughness sections in the summer bed. This range of parameters gives a sphere of fitting for the used surrogated model, namely a NARX. The NARX is trained and verified on the different potential runs of D-Flow FM. With this a highly accurate trained NARX is created that has an R^2 between 0.99 and 0.75, for the best FM run and the worst FM run respectively. This means that the NARX can mimic the D-Flow FM model.

The NARX is used in combination with an interior point barrier function algorithm to reconstruct the original discharges and roughness values of the summer bed. The resulting original discharges however were unphysical. Usually the model resulted in discharges over $4*10^4$, which is extremely more than any literature. This shows that the method has some flaws in this specific case.

The flaws are most likely caused by one of the following three reasons: too many variables, outside the sphere of fitting, or failing optimization algorithm. First, the too many variables are mostly the number of roughness sections compared to the number of measurement stations. There are relatively few measuring stations in the area, compared to the number of roughness sections. This could be solved by reducing the number of roughness sections. Second, being outside the sphere of fitting, means that the NARX cannot represent the physical model for the measured water levels. This can be fixed by changing the type of experiments used in the training phase. Third, the failing of the optimization algorithm, this means that the problem does not fit to the requirements of the optimization algorithm. To solve this another optimization algorithm can be used.

Concluding, the accurate representation of the physical models shows that the NARX is capable in representing the physical model. This is possible even though there are changes to the physical properties of the area. However, it showed that in the current approach there are still difficulties to reach a desired water level. This means that further research is needed in the three aspects that are most likely causing this problem.

Table of contents

1 Introduction

1.1 Background

The flood security policy in the Netherlands has changed over the last years. This change gives a move from probability of exceedance to flooding probability (Dutch Ministry of Infrastructure and the Environment & Ministry of Economic Affairs, 2014). The probability of exceedance is the chance that a water level is exceeded, whereas the flood probability is the chance of actual flooding. This change led to a water safety question for most of the river dikes.

Dikes are designed according to conservative methods, since design discharges are uncertain. This means that the most extreme situation is used, leading to excessively strong designed dikes. With the new policy the safety standards have increased once again, leading to more expensive reinforcements. With better methods to determine the design discharges, the uncertainty in these can be reduced. However, this requires more detailed models to describe the water systems. With these models, the uncertainty in design discharges can drop leading to potential smaller dikes and less expensive solutions. Most of this uncertainty is caused by the high return periods, whilst having only relative short time series of data to base the extreme events on. To extend this time series, more events of the past could be added by creating historic flood reconstructions. For the reconstructions of floods detailed two dimensional models are required in the embanked areas, as the directions of the flow in the flooded area is unknown (Hesselink et al., 2003).

Two dimensional models require more computational power than the simpler 1D models. As the computational power is limited, this problem can be addressed by using a metamodel (Behzad et al., 2009; Simpson et al., 2001). Metamodels are a simplified version of the original model, which are suitable for a specific application (Kolkman et al., 2005). These metamodels, also called surrogated models, are a simplification of the original model. These models are categorised in two types: lower fidelity modelling and data modelling (Razavi et al., 2012a). Lower fidelity modelling works by reducing the number of physical processes included (Razavi et al., 2012a). Some options for lower fidelity modelling are: creating a coarser grid, moving from 2D to 1D modelling, and reducing the temporal resolution. Data modelling works by fitting a relation between the input and output of the original model (Razavi et al., 2012a). Some examples of data modelling are; the interpolation of different measuring stations, the use of polynomial functions to receive timeseries, and the use of more advance technique like neural networks. The potential time gain for each individual model run is larger for data models than for lower fidelity modelling as it removes the physical limitations entirely.

With such a metamodel the original model is mimicked. This means that the metamodel is usable for the same application as the original model within the sphere of fitting (Razavi et al., 2012b). This sphere of fitting is determined by the data of the original model. The more data of the original model is available the larger the sphere of fitting. The use of these surrogated models in operational sense has shown to be effective in speeding up the simulation whilst remaining accurate (Duong et al., 2018; Matta et al., 2018). However, once the inputs are outside of the sphere of fitting the results of these models are no longer valid. As these metamodels have lost all physical properties extrapolation of the input is not possible. The metamodels will give a result even though the input is outside the sphere of fitting, however the error of this result remains unknown.

Various model inputs are missing for historic reconstructions. These inputs are things like discharge and roughness values, which means that physical modelling becomes troublesome. Whilst the calculation times do not allow for iteration of these parameters, with 2D models this problem becomes even larger. To tackle these problems metamodels could bring a solution. Whilst the

metamodels have lower calculation times, the correct input out of a set of inputs might be determined. For the case of 1809 that is used in this study, both the roughness and discharge wave are unknown. Therefore, the use of a surrogated model to fit these inputs to the water levels is explored. This should give the benefit that it is possible to remake the discharge wave towards the actual measured water levels.

1.2 Objectives

Till now surrogated models are used for the speedup of operational models and design optimizations (Simpson et al., 2001). In this study the goal is to explore the potential of using these surrogated models for reconstructing discharge wave of 1809 in the Rhine system.

1.3 Scope

The focus in the study is placed on data modelling techniques. This means that the lower fidelity techniques are excluded. Lower fidelity models are based on the same core physical processes as the high-fidelity models. Therefore the results become less accurate in exchange for more speed (Razavi et al., 2012b). While the data based surrogated models have their calculation efficiency gains from using different methods. This means that the calculations that a data model makes are relatively easy and simple (Cheng & Titterington, 1994). Whereas most physical models have difficult routines using multiple differential equations and solutions for these, data models consist out of linear algebra. Classical computing is extremely optimized for linear algebra, leading to a large calculation time gain.

1.4 Report outline

This thesis report will first explain the methodology of the study. This can be found in Chapter 2. After this the different results will be discussed, first the accuracy of the surrogated model, in chapter 3 and the discharge reconstruction in chapter 4. With furthermore the discussion, conclusion and recommendations in the follow up chapters.
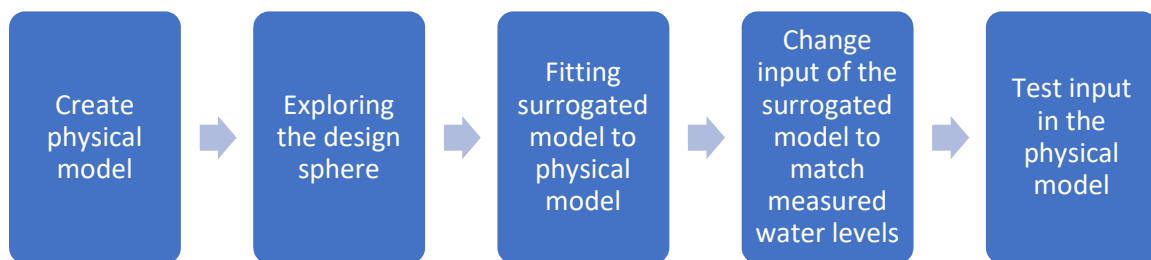

Figure 1: It gives an overview of the overall methodology of this thesis, with an iterative process

2   1809 flood

During the winter of 1809 large areas in the Netherlands were flooded. It is currently still unknown what the discharges were during this flood. In total 100,000 people were affected by the floods in 1809. Of these 275 people were killed and in total 2,000-3,000 equines, bovines, and pigs were killed and about 1000 homes were completely destroyed (Driessen, 1994). The floods were caused by the cold period before the first flood event that occurred around the 10th of January 1809. A second flood event followed around the 25th of January flooding more regions downstream.

The focus is placed at the upstream flooding event in this study, meaning the first one. This choice has been made to reduce the unknown variables in the study. For the second flood event, the Meuse river has met with the Rhine delta. This means that both the discharge of the Rhine as well as the Meuse effect the water levels in the system. This would become to complicated for the given case and methodology, as such the case is reduced to only the first flood event, and spatially cut of before the Meuse becomes a part of the system.

From half December 1808 till the 10th of January 1809 the Netherlands had a period of heavy frost (Lintsen, 2009). Due to the low temperatures, the rivers (Rhine, Meuse) were completely frozen. Around the 10th of January the temperature rose leading to melting ice and the creation of drift ice. These ice piles hit each other chaotically, leading to a large iceberg at the south of Arnhem. This iceberg blocks the Nederrijn river leading to a larger discharge in the already full Waal river, leading to dike failures in the area (Driessen, 1994; Lintsen, 2009).

2.1   Study area

The study area is a part of the Rhine branches in the Netherlands. This consists of the part located between Emmerich and Tiel/Zupthen. In the study area, around the year 1800, some measurement stations are available. These stations are spread over the branches, but in a low spatial resolution. The stations measure the water levels at several locations, namely: Arnhem, Nijmegen, Pannerden, and Doesburg. The temporal resolution of the stations is daily. Using maps of 1810-1850 the geophysical properties of the area are extracted. This is explained in more detail in paragraph 2.2. Figure 2 gives an overview of the study area with the locations of the measuring stations and the river courses of 19th century placed on the current day topography. Some difference exists between these two situations. The largest visible deviations are the course from Pannerden to Doesburg, which is in the current day less meandering then in 1800. Another visible deviation is downstream of Doesburg, where a meander is cut-off in the current day.
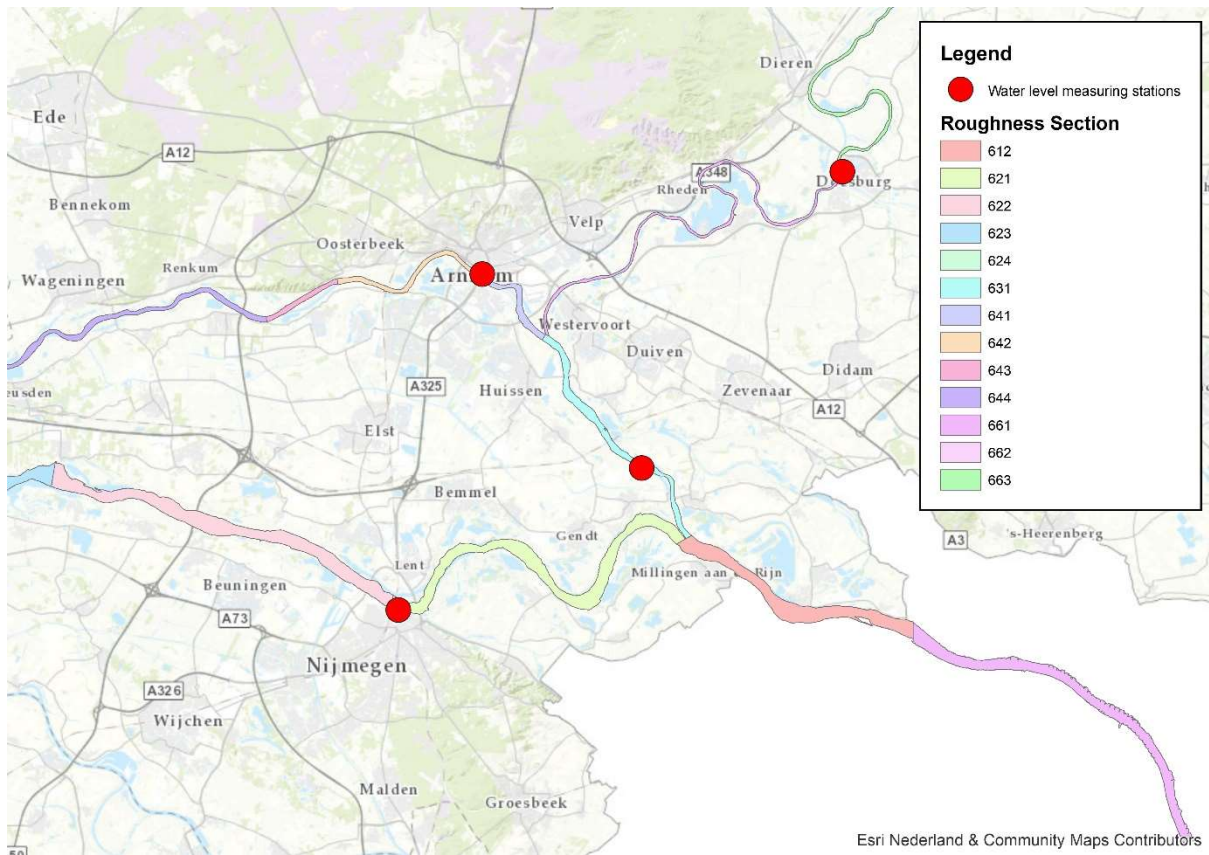
Figure 2: An overview of the roughness sections and measuring stations in the study area

2.2 Reconstructing the historic setting

As the case in this study is a historic setting, the situation first needs to be reconstructed from several historic sources. This is done by combing information into a physical model of the 1809 flood, more over this in Chapter 3. For such a physical model some inputs are required; things like the location of the summer bed, the location of the floodplains, the dikes, the height of the area, the flow roughness. These inputs can be extracted from historic information, this chapter will give an overview of the reconstruction and how the information is gathered.

2.2.1 Data sources for reconstruction

Each input for the physical model has one or more data sources, for the reconstruction. Each of these inputs will be discussed in more detail, how the reconstruction occurred. Table 1 contains an overview of the different maps and datasets used to reconstruct the 1809 flood. For each the name of the source is given, together with the year and author.

Table 1: Overview of the data sources used for the reconstruction

Model inputs	Data extracted from	Time data source	Source
Roughness values embanked areas	Historic land use 1900	1900	(Alterra, 2004)
Roughness values summer bed	Theoretical manning values	-	(Chow, 1959)
Height in floodplains and embanked areas	AHN	1850	
	Baseline	2000	

Dike locations	Register der Peilingen	1830-1850	(Nederlandsch Aardrijkskundig Genootschap, 1837)
	Dike breach Oosterholt	1810	(Ewijk, 1809a)
	Dike breach Loenen	1810	(Ewijk, 1809b)
	Dike breach Loo	1810	(Ewijk, 1809c)
Dike heights	Waterstaat Kaarten 1st generatie	1872	
Bathymetry of the summer bed	Tables of Register der Peilingen	1830-1850	(Nederlandsch Aardrijkskundig Genootschap, 1837)
Summer bed course	Register der Peilingen	1830-1850	(Nederlandsch Aardrijkskundig Genootschap, 1837)
Water levels	Rijkswaterstaat waterinfo	1809	(Rijkswaterstaat, n.d.)

2.2.2   Roughness embanked areas

The roughness in the flood plains and embanked areas are extracted using the land use of 1900. The land use of 1900 is used, as this was a readily available digital map with no significant changes to 1809 in the land use. This map is a digital colour map of Alterra (2004), each colour represents a type of land use. Using the classification learner of Arcmap all the available colours are categorised using the legend of the original map. With this information the map is transformed in polygons with land use categories attached to them. These polygons are later used, to transfer roughness values in the physical model.

2.2.3   Roughness summer bed

The summer bed roughness values are determined using theoretical manning values. For this the manning values for different type of river segments are combined. This is done using Chow (Chow, 1959), these theoretical parts are added together. While all river types never have an absolute single manning value, this leads to a band of potential manning values. For the experiments later, these values are varied within this unknown bandwidth of manning values.

2.2.4   Height of the floodplains and embanked areas

The terrain heights except for the summer bed are generated using a height map of the Netherlands from 1850. The resolution of this map is 250m by 250m. This data gives all the winter beds, and inner dike area heights within the Dutch part of the study area. Next to this Dutch section, the German inner dike area and winter bed requires height information. These heights are created, using the current day heights for the embanked area and for the winter beds a height correction is added on top of the current day heights. This height correction is the average difference between the current day heights and the Dutch 1850 heights in a segment of winter bed close to the border, but where both datasets are available. The section used is the north side of the Rhine from Pannerden to Lobith. Figure 3 provides an overview of the different location where this height data is used. No significant changes in height are noted between 1809 and the published set in 1850, as such the 1850 data is used for 1809. This height data is available in the Dutch part of the study area. For the German part the height data of 2000 is used. This height data is manipulated in the floodplains and summer bed, to match the 1850 dataset. For this a section of the Dutch part just past Lobith is used to calculate the average deviation between these datasets. The embanked areas in the German area are just determined by the 2000 dataset.
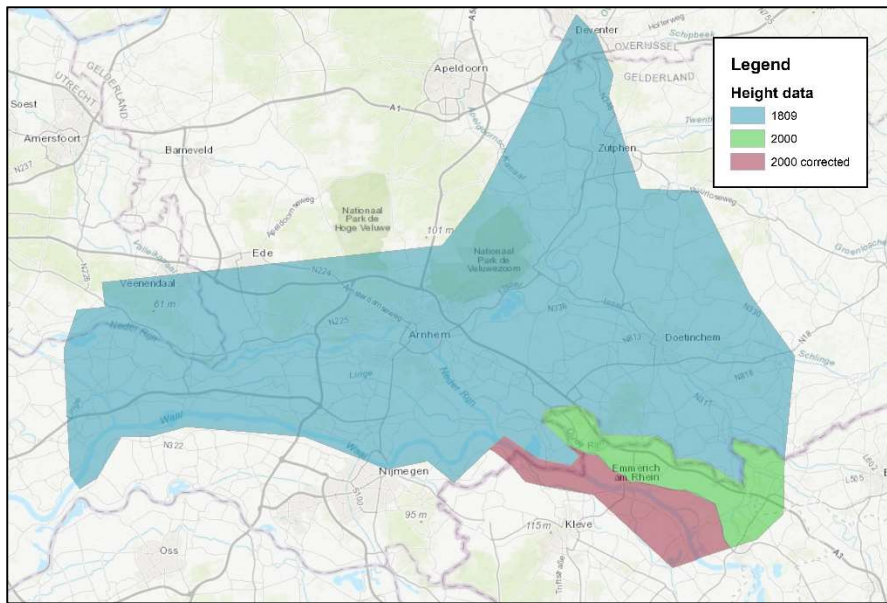
Figure 3: Used height data for the D-Flow FM situation

2.2.5   Dikes

The dikes in the Netherlands have changed in the last two centuries. To have the outer dikes as accurately as possible in the model, the location has been extracted from old river maps of 1830-1850, with corrections of the dike breaches in 1809. Paragraph 2.2.5.1 describes this extraction in more detail. Next to the location also the heights are needed. These are extracted from the "Waterstaatkaarten" of 1875. Around the dike breaches more detailed maps were available that are used for the dike heights there.

2.2.5.1   Location

The location of the dikes is digitized from the 1830-1850 maps in 'Register der Peilingen' (Nederlandsch Aardrijkskundig Genootschap, 1837), together with detailed views of the dike breaches ("Beeldbank (Gelders Archief) - Gelders Archief," n.d.). With the dike breach maps a more accurate representation of the outer dike location is made. The river reconstruction map is from 1810 after this flood event, as such the dikes are relocated due to the dike breach. Using the river map of 1810, the location of the dikes is digitized by creating line elements following each of the outer dikes. All small embankments in the winter bed are ignored in this research. Next to the location the dike breach maps also have information about the size of the breach leading to a more accurate representation of these breaches in the physical model. However, some of the maps have become less opaque over time, leading to uncertainties in the exact dimensions. These width inaccuracies are in the order of tens of meters, leading to relatively small uncertainty. As the grid for the physical model will not be smaller than tens of meters. The physical model will map all dikes to the grid, so the breach will always be a natural number of grid cells wide.

2.2.5.2   Dimensions

The dimension of the dikes, which is modelled as only the height is extracted using the Waterstaatkaarten. These maps contain the dike system of 1872, with the corresponding height at several locations. The height data is placed as points on the locations measured on the different dike segments. After which for every vertices of the original dike line the heights are interpolated linearly. This linear interpolation is done with the length of the dike segments, as such a 1D interpolation over the dike.

2.2.6 Course of the river summer bed

In cooperation with the university of Utrecht, a reconstruction of the river summer bed is created. This is done by combining the old maps of 1830-1850 which are part of the Register der peilingen (Nederlandsch Aardrijkskundig Genootschap, 1837). These maps have been spatially placed in the correct location, using multiple recognition points through Bas van der Meulen (UU). With this information the entire course of the summer bed is digitised. This information is the basis for the bathymetry of the Rhine river delta. More information on the interpolation method can be found in the next section.

2.2.7 Bathymetry of the summer bed

From measurements every 1 km of the river the bathymetry of the entire river had to be reconstructed. Caviedes-Voullième et al (2014) discusses the possibility of a cubic hermit spline to interpolate the bathymetry over the course of the river. This method works by using a cubic Hermite spline function to determine the thalweg. Cubic Hermite spline is an interpolation method that requires both the function to be continuous and differentiable in the data points. The method of Caviedes-Voullième et al (2014) uses the spline through the thalweg, with perpendicular cross sectional sections on top. These cross-sectional sections are divided in several points. The points are then connected from cross section to cross section, on which linear interpolation is used. This method only works in symmetric rivers, if the thalweg is in the centre. In asymmetrical rivers the resolution of the interpolation gets skewed in the river. As the river branches in question have a lot of curves and moving thalwegs over the cross-sectional direction, this method is not applicable, therefore an extension of this method is used.

The more advanced method is using the thalweg generated by the cubic Hermite spline to generate a curvilinear grid in line with the river. Figure 4 gives an overview of this curvilinear grid that is created shown in the red lines. The method uses this grid to give every sample (coloured points) an x,y coordinate in this grid. A grid for the interpolation results in then created between the summer bed boundaries and in a fine spatial resolution around 2-5m. These points are filled by linearly interpolation in the curvilinear grid. The small yellow dots in the black lines are the locations of these points in Figure 4. Due to the zoom level and the grid size, does this fall behind most of the lines. The interpolated values will follow the course of the river, as the curvilinear grid values the direction in this manner.
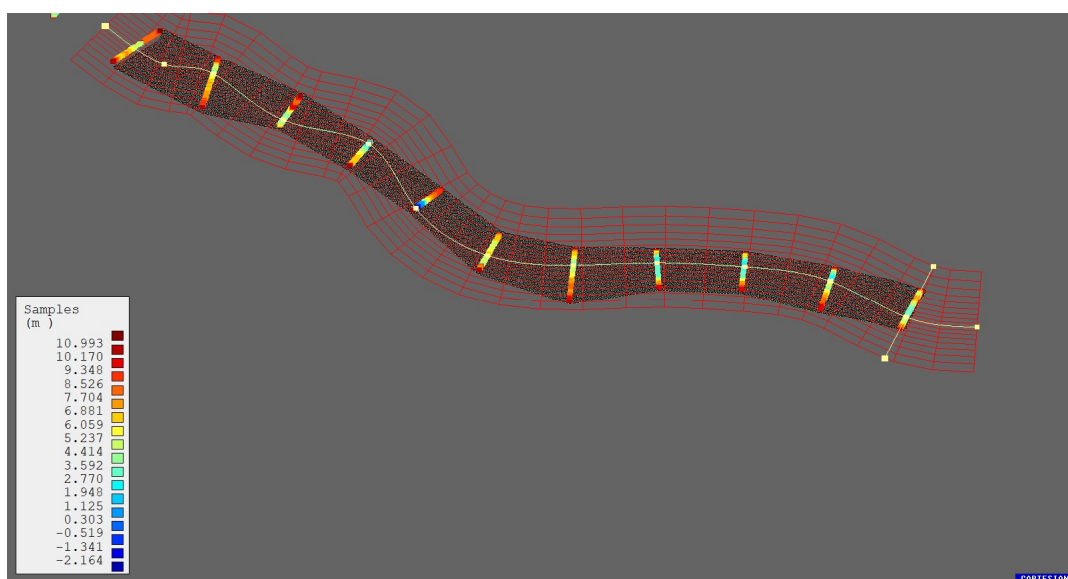

Figure 4: The Curvilinear grid interpolation

2.3 Discussion of the reconstruction

The reconstruction of the historic case is just a part of this thesis, as such some assumptions have been made to make the reconstruction possible in the given timeframe. These assumptions reduce the accuracy of the reconstruction. However, for the given research aims of this thesis the assumption will have a small effect. When the case of 1809 reconstruction is the main goal the following aspects should be fixed.

First, the biggest error is caused by missing height data in several sections of the study area, especially in the German locations at the upstream section of the study area. Due to the limited size of the section being a few kilometres, the effects on the results remain limited.

Second, the location of dikes and the heights do not match on the maps (Waterstaatkaarten and the Register der Peilingen). This can be caused by projection issues, or different georeferencing locations as the maps are part of two different datasets. Whilst the shape of the dikes was the same on both maps, the likeliness that the heights are within a couple of meters on the correct locations is very high. Therefore, the effects of this potential inaccuracies are small.

Third, the interpolation of the summer bed, is created from cross sectional data every kilometre. This means that the bathymetry course is not located in the input data, as such the river course could have been different instead of following a perfect spline. Whilst no more detailed data is available this inaccuracy is to be accepted. The eventual effects of this assumption will be limited. As the main flow direction is ensured using the thalweg spline (Caviedes-Voullième et al., 2014), which is the base of the curvilinear grid.

Fourth, the roughness of the summer bed is determined using the physical roughness values. However, in this roughness a model grid roughness should be included. Due to the model grid numerical inaccuracies start to occur, these are usually caught by the roughness as the model is calibrated using roughness values. These values are missing, which could potential lead to wrong results. As the summer bed roughness and the discharge are unknown values, calibrating these was not an option, till the final stage using the surrogated models.

Final, the topographic map used for the land use in the embanked areas is a lot more recent than the event. This gives a discrepancy of about 90 years, as such some of the land uses have changed, mostly that the cities were increased somewhat in size. However, no major changes were spotted between the maps, together with the relatively small sensitivity of the model to these values (Bomers et al., 2018a), means that this error is neglectable.

2.4 Measured water levels during the flood

Using the measurement stations that are located at Arnhem, Doesburg, Nijmegen and Pannerden daily water levels are available. These water levels will be used to reconstruct the original discharge wave in this situation. These measuring stations give the water levels on a temporal resolution of every day; however, the exact time of measurement is unknown. Figure 5 shows the measured values. It shows different behaviours between the measurement station. In this ice dam effects are present, with steep discharge peaks.
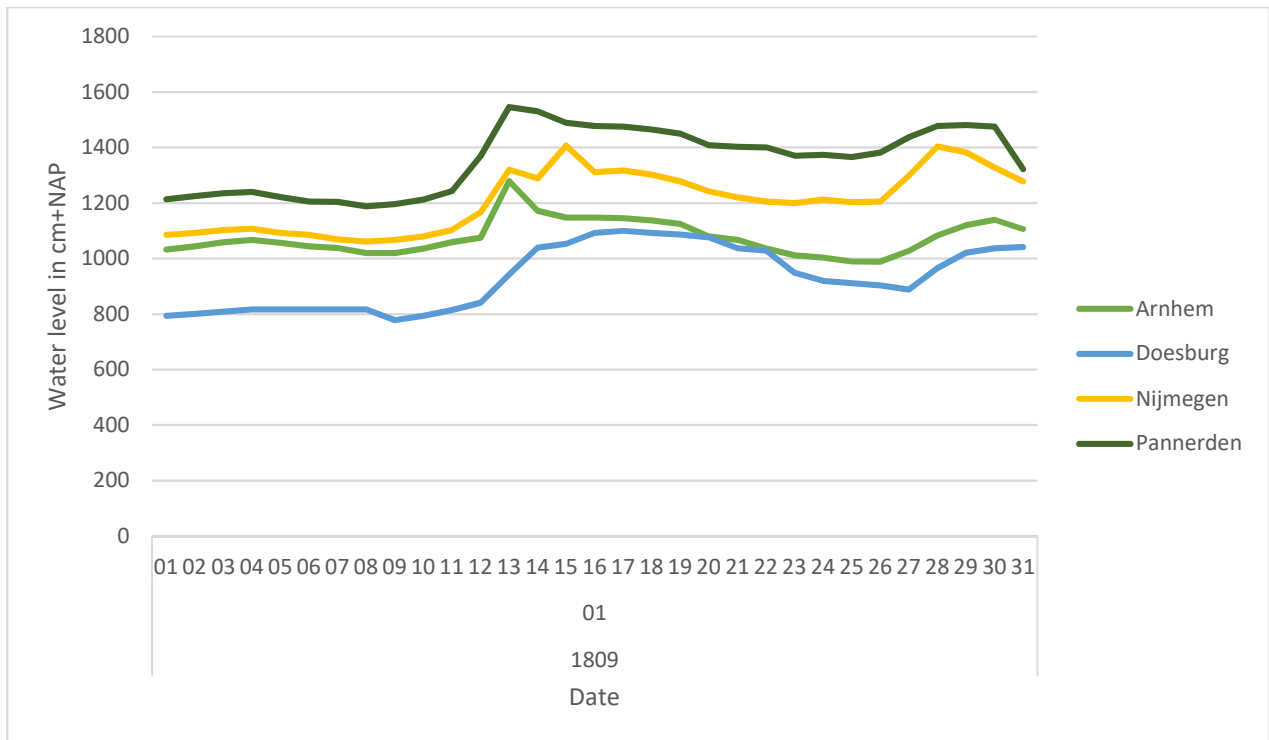
Figure 5: Measered waterlevels 1809

3 Physical model

In this thesis a physical model is created of the 1809 case, this physical model is a bunch of hydrodynamic equations. The hydrodynamic modelling software used in this thesis is D-Flow FM, also known as flexible mesh (Deltares, 2016b, 2016a). The advantages of this software are the ability to combine different types of grid shapes and sizes. This allows the solution to become more accurate in interested areas and faster overall. Physical functions are complicated to calculate, as most of the functions use partial differential equations. These kinds of equations are solved using numerical schemes. For this solving method a few things are needed. First, the numerical schemes require a grid for the study area. Second, the boundary conditions must be determined. Third the roughness in the study area must be included.

3.1 Model grid

Using flexible mesh, a smooth and orthogonal grid is required (Deltares, 2016a). This means that the neighbouring grid cells should be about the same size and the centrelines should be perpendicular to the cell boundaries (Deltares, 2016b). Bomers et al (2018b) mention that hybrid grids are better capable of determining realistic flow velocity patterns as long as the grid resolution does not change quickly between the different parts of the grid. The advantage of D-Flow FM is that such a hybrid grid can be made using different grid sizes and shapes over the model area if the differences in sizes are not too large. D-Flow FM is used with a curvilinear grid in the summer and a triangular grid in the floodplains and embanked areas. The resolution of the triangular grid is decreased towards the embanked areas from the boundary of the summer bed. This is done as the interesting results and spatial differences further from the summer area becomes sparse. The eventual inner triangle grid has a resolution of 350m sides.
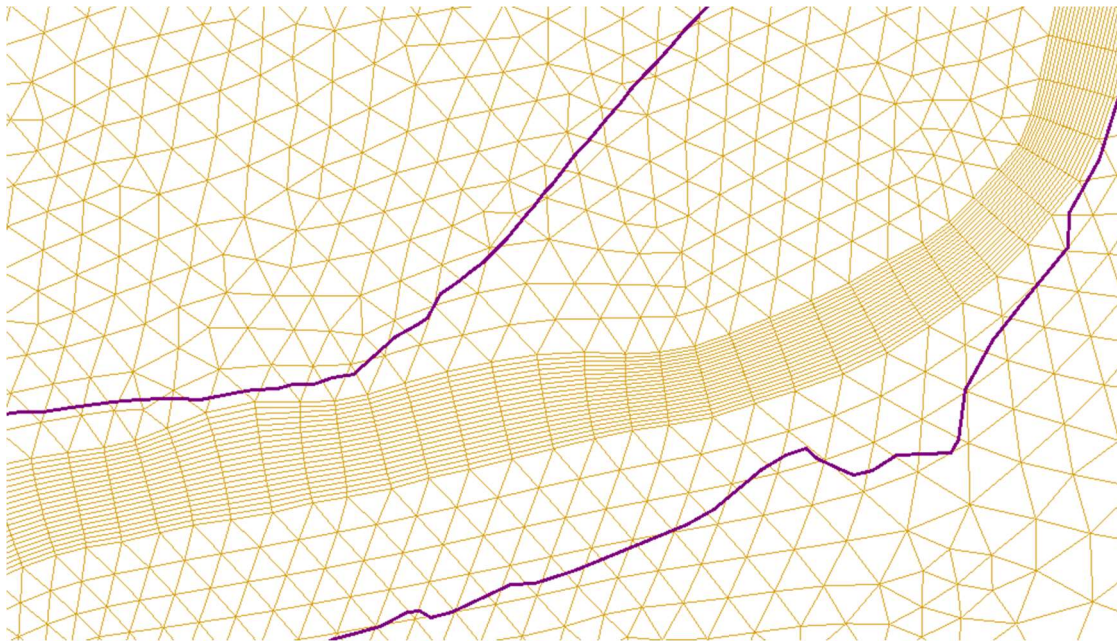
Figure 6: An example of a piece of grid, in which the dikes are located with the purple lines. A curvilinear grid in the summer bed, and triangles in the floodplains
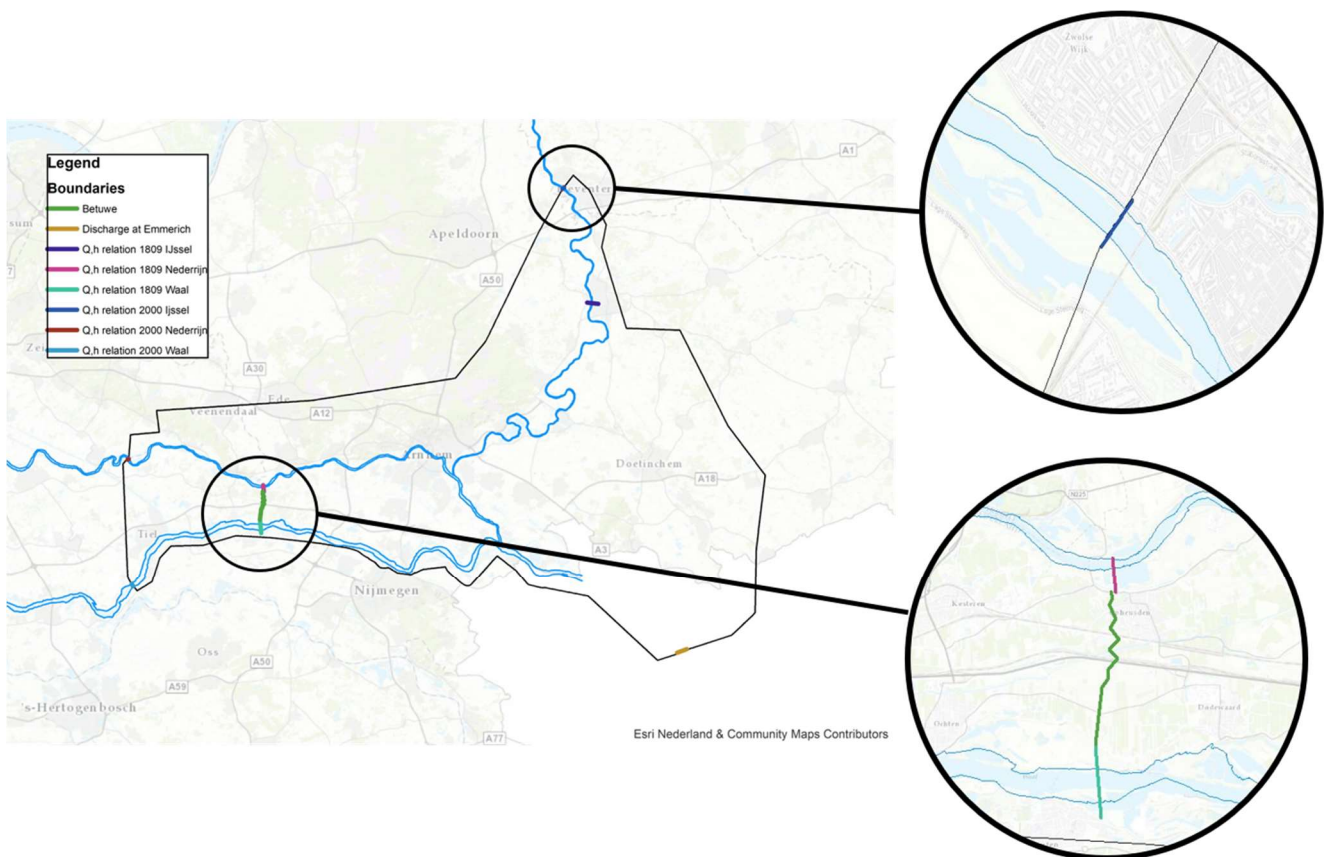
Figure 7: Overview of the D-Flow FM boundaries

3.2   Model boundary conditions

The equations, which are the basis of this software require boundary conditions. In total there are five different boundary condition types used. First, that no water can flow over the boundary, this type is used most often for all segments in which water flow is not directly present. Second an

16

amount of flow over time is used, this is used at the upstream boundary and just shows the discharge over time. Third, Q-h relations are used, these boundary conditions couple the discharge at the boundary to the water depth that should be present there at the same time. Combining these conditions means that the equations of this model can be solved.

The first type of boundary conditions is located on the black lines of Figure 7, these are the boundaries of the grid. This is just to prevent the water from leaving the model at unexpected locations. The other colour lines in Figure 7 show the Q-h relations at several locations and the discharge at Emmerich, which places a discharge wave on the location. This discharge wave is created using a theoretical shape. This decision has been made to ensure the possibility of testing the new method, whilst remaining having a natural shape. The shape of the discharge wave is shown in Figure 9, this consists out of a baseflow of 6000 m^3/s, in a combination with a flood wave in a gaussian curve. The baseflow is chosen as the maximum flow in the Rhine river with a return period of one year (Rijkswaterstaat, 1994). This reduces the warm up period as all the river branches are filled with water including floodplains. The gaussian curve ensures a single peak in a smooth fashion. This smoothness is useful as it reduces the problems that later functions can bring.

For the Q-h relations a different approach is used, the river boundaries are determined using the current day Q-h relations further downstream. As the differences between 1800 and the current day are not huge, a boundary 20 km further downstream suffices. This is far enough that there is no water staying on the river, which makes the locations of the correct boundaries not be in steady state. Combining all steady state water levels at multiple constant discharges, namely 4000, 8000, 12000, 16000, and 20000 m^3/s leads to a Q-h relation. This Q-h relation is made for each river boundary and shown in Figure 8. The betuwe boundary has been determined differently. As the dike breach gap is known, the flow from there over the land will be the same independent of the discharge in the river. This means that using the maximum discharge and then considering the Q-h relation on the land will give a nice overview of the flow on this border. This is determined by creating a storage hole a bit downstream, which is filled. When the storage is full, the Q-h relation is cut. This means that the relation assumes infinite storage downstream. As the period for the calculation is limited this is fine for the scope. The extend in which calculations will be done is from the 8th of January till the 20th of January.
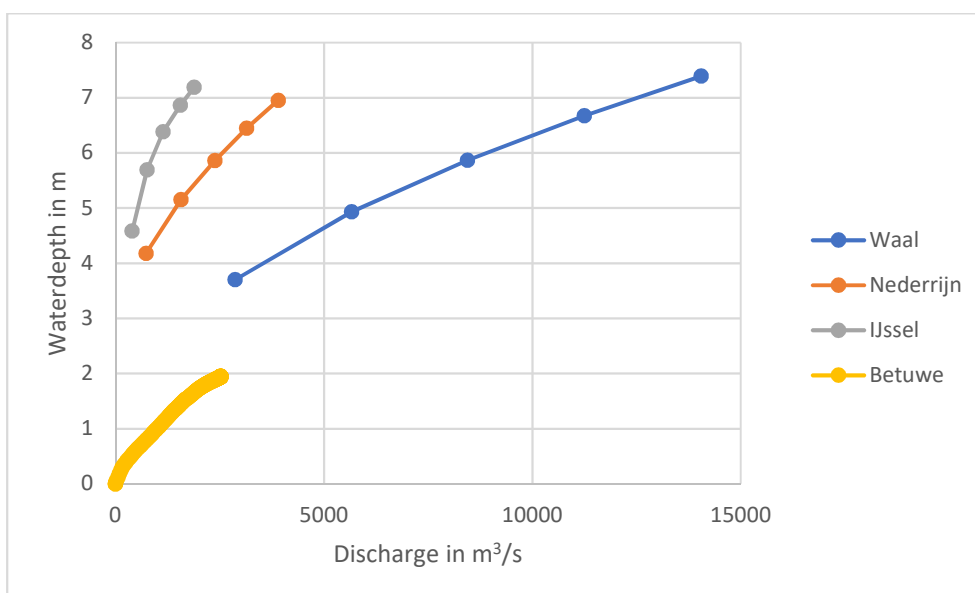
Figure 8: The Q-h relations as determined on the boundaries, all of these are for 1809

3.3 Dike breaches

In 1809 there are four major dike breaches that have been included in this study. All these dikes have been breached between the 12th of January and 15th of January. Three of the four dike breaches have detailed maps of the location and size of the breach. For the fourth dike breach, the size has been determined using a dataset containing all major dike breaches of the last century in the Netherlands. The size is then the average data of all breaches in this dataset. Figure 10 gives an overview of these location, of which A has been ignored due to the very small embanked area. The location of the fourth dike breach is taken from the global map, shown in Figure 10.
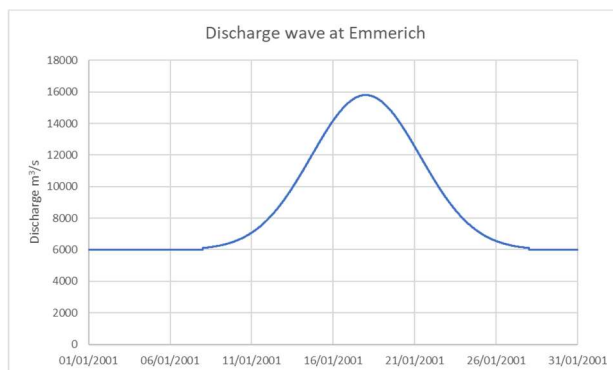
Figure 9: Upstream boundary condition
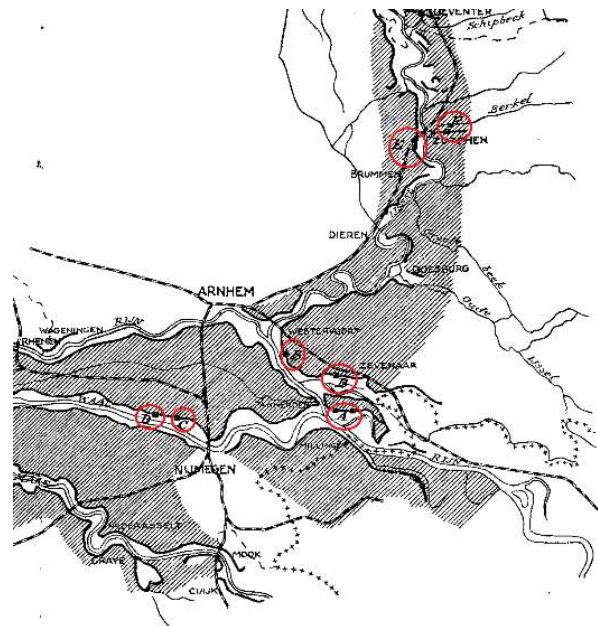
Figure 10: Location of the dike breaches, the bottom five have been considered, in which A is ignored due to the small area that could flood, B is combined into the bottom breach (Dutch Ministry of Infrastructure and the Environment, 1926)

4 Surrogated models

Current day modelling tries to become closer to reality every year and requires more computational power. Reducing the computational power, surrogated models can be used. These models represent the model but in a simpler form and so less calculation intensive form. The disadvantages of surrogated models are the slight loss of accuracy with respect to the original model (Razavi et al., 2012b). Within surrogated modelling, there are two types of models, namely response surface modelling and lower-fidelity modelling. Response surface modelling use data-driven function approximation techniques to empirically approximate the original model response. These types are also known as metamodels or proxy models as it is a model of the model. Lower fidelity models are physically based simulation models but less-detailed compared to original models (Razavi et al., 2012b). In this thesis the focus is placed at response surface modelling. The setup of these models is discussed in the following paragraphs, with first the experiment design, and afterwards the different surrogated models that will be used.

In this thesis surrogated models are used to get an estimation of the discharge wave that occurred in 1809 during the first flooding. It is impossible to achieve this result with the physical 2D models as the calculation requirements are too large. Figure 11 provides an overview of the different steps that are taken in this thesis. The first two boxes are discussed in the previous chapter of the

methodology, this chapter focusses on the third to the final box. With the physical model a surrogated (data) model is setup for pre-determined experiments, these are discussed in the next paragraph. Using these experiments, the surrogated model is fitted, this is discussed in paragraph 4.2. With this surrogated model the input is changed to fit the water levels measured in 1809.

```
┌─────────────────┐   ┌─────────────────┐   ┌─────────────────┐   ┌─────────────────┐
│ 1809 geophysical│──▶│ Flexible mesh   │──▶│ Surrogated model│──▶│ Fit the         │
│ data collection │   │ model of the    │   │ of the Flexible │   │ surrogated      │
│                 │   │ 1809            │   │ mesh model      │   │ model to the    │
│                 │   │ geophysical     │   │                 │   │ measured 1809   │
│                 │   │ location        │   │                 │   │ data            │
└─────────────────┘   └─────────────────┘   └─────────────────┘   └─────────────────┘
```
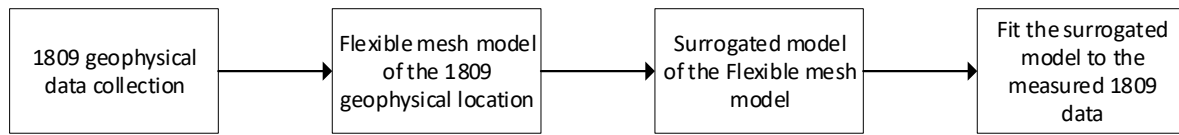
Figure 11: Overview of the surrogated method to get 1809 information

The goal of the surrogated model is to allow for the reconstruction of the original discharge wave in 1809. To ensure this possibility the surrogated model should fit the D-Flow FM data, as this allows it to get the relation between the input and the water levels. To ensure this fitting an experiment design is built. An experiment design means a list of input parameters. These input parameters are chosen such that the entire options of potential inputs is covered (Saltelli et al., 2008).

4.1   Experiment design

The quality of an experiment design depends on two features of the dataset, namely the bias and variance. The bias of a dataset is the quantification of the extent to which the surrogated model outputs differ from the true values calculated as an average over all possible data sets (Queipo et al., 2005). The variance of a dataset is the extent to which the surrogated model is sensitive to a particular dataset. Each dataset corresponds to a random sample of the function of interest (Queipo et al., 2005). The bias can be reduced by including more data points, however the variance is improved by smoothing the data (Queipo et al., 2005). Therefore, reducing the bias will increase the variance and vice versa. For an experiment design based on deterministic computer simulations the focus lies on bias reduction (Queipo et al., 2005). Without knowing the objective function of the surrogated model, the best practice is to use an uniform sampling method as this keeps both the variance and the bias low (Koziel et al., 2011). For the creation of experiment designs many methods exists, out of which only Latin Hypercube Sampling (LHS) and Orthogonal Arrays (OA) will sample uniformly (Queipo et al., 2005). Therefore, usually either of these strategies or a combination of both is used. LHS is used in this thesis as it has advantages over OA, in being more flexible (Koziel et al., 2011). LHS works by stratified sampling. This means that the distribution is split up in sections of equal probability. In each of these sections a sample point is taken randomly. This means that the distance between samples could vary slightly but will always be limited by the sections. The section in LHS are called the level of the LHS method. The number of levels of the LHS is a trade-off between computational expense and the accurate representation of the distributions (Bomers et al., 2018a). Therefore, it has been chosen to take 8 levels as these follow the distribution and this leads to an allowable computational expense. In total four 8 level designs have been made to allow for enough training and validation data. Two sets of this design can be found in Table 2 in which each number tells which stratified section the value should be placed in.

Table 2: The LHS experiment design

	Roughness sections														Dike	Discharge
Run	**611**	**612**	**621**	**622**	**623**	**624**	**631**	**641**	**642**	**643**	**644**	**661**	**662**	**663**	**Time**	**Peak**
1	0	6	3	6	1	0	5	2	1	7	4	7	4	4	4	6
2	4	2	1	0	5	2	6	6	0	6	0	6	2	5	6	3
3	5	7	4	4	0	4	0	5	2	4	6	5	7	7	5	4
4	2	0	2	5	7	7	4	4	3	0	7	3	0	0	0	5
5	3	1	5	1	6	1	3	3	6	1	1	0	5	1	3	7
6	6	5	0	2	3	3	1	1	7	3	3	1	1	2	1	0
7	7	3	7	7	4	6	7	7	4	5	2	2	6	6	2	1
8	1	4	6	3	2	5	2	0	5	2	5	4	3	3	7	2
9	4	6	5	7	2	6	5	1	7	4	6	3	4	1	5	2
10	6	7	2	4	0	0	4	4	2	0	4	6	2	0	2	7
11	1	1	7	2	1	2	3	7	4	6	5	0	0	5	4	6
12	5	2	3	1	3	4	6	3	5	7	3	1	6	7	1	5
13	0	5	1	3	6	3	7	6	1	3	2	2	5	6	7	3
14	3	0	0	6	5	5	0	2	6	5	1	4	3	2	3	4
15	7	4	4	0	4	7	1	0	3	1	7	5	7	3	0	1
16	2	3	6	5	7	1	2	5	0	2	0	7	1	4	6	0

Table 2 gives a change for the roughness sections, this means that section as shown in Figure 2 changes the manning coefficient as determined within the stratified sample. Next to these roughness values, the dike breach time is changed. Finally, the peak discharge is changed, this is done by scaling the discharge wave towards the new peak. Therefore, all the discharges during this wave will change. The shape of the wave as presented in Figure 8 will remain the same however.

For each of the runs a random value in the segment of the LHS is chosen. The random values are created from different distributions for each parameter. Table 3 presents an overview of these distributions, by naming the distribution and the minimum, mode, and maximum values of each parameter. The sections for LHS are divided uniformly of size 0.125, meaning that a 0 in Table 2 corresponds to a random value between 0-0.125. Using the distributions and these sections the values for the parameters are created. For the roughness values a beta distribution is chosen, as this gives a bounded distribution with a mode value. This is useful when a mode, min, max is known.

Table 3: Overview of parameter ranges

Parameter	Min	Mode	Max	Distribution	α	β
611	0.03	0.036	0.044	Beta	2	2
612	0.03	0.035	0.044	Beta	2	2
621	0.0345	0.0391	0.0506	Beta	2	2
622	0.035	0.037	0.049	Beta	2	2
623	0.035	0.037	0.045	Beta	2	2
624	0.035	0.035	0.045	Beta	2	2
631	0.025	0.029	0.039	Beta	2	2
641	0.02875	0.0322	0.04025	Beta	2	2
642	0.04025	0.04255	0.05175	Beta	2	2
643	0.035	0.035	0.04	Beta	2	2
644	0.02875	0.0322	0.04025	Beta	2	2
661	0.045	0.048	0.06	Beta	2	2
662	0.0345	0.0368	0.046	Beta	2	2
663	0.025	0.025	0.03	Beta	2	2
Dike breach Time	0	-	24	Uniform	-	
Peak discharge	7000	-	18000	Uniform	-	

4.2 Artificial Neural Networks (ANN)

The surrogated model used in this thesis is an artificial neural network (ANN). ANNs have two main fields of operating, namely classification and regression. In this thesis, only the regression aspects are interesting, as the goal is to use it as a surrogated model of an original model. ANNs have been used for complex problems, like face, speech, handwriting recognition; currency exchange rate prediction; chemical processes optimization; cancer cell identification; and spacecraft trajectory prediction (Cheng & Titterington, 1994). Within Civil Engineering the uses of ANNs have been in emulators for reservoir operations problems, for instance to capture patterns between flow rates and storage levels (Raman & Chandramouli, 1996).

The type of ANN that is used in this thesis is a nonlinear autoregressive network with exogenous inputs (NARX). The choice for a NARX is made as these models are meant for regression modelling of time series. The discharge water level relation that is required, is nothing else then such a time series (Hagan et al., 1995a). A NARX exists out of 3 layers, namely the input, a hidden layer, and an output layer. Figure 12 gives an overview of a NARX in a general layout. The version used in this thesis has a feedback loop, meaning that next to the external inputs ($\mathbf{p^1}$ in Figure 12) the output is looped back. The effect of this feedback is that the previous water level effects the next time step, meaning that impossible physical changes cannot occur due to persistence in the system, e.g. quickly changing water levels. Each of the layer consists out of several nodes, also called neurons. In each neuron a scoring function is present, that considers the input of all previous nodes times weights and determines its output from there (Hagan et al., 1995b).
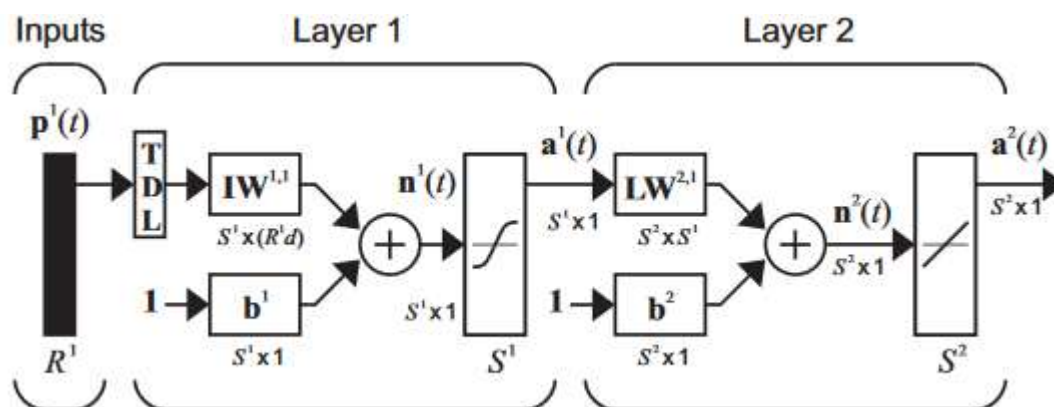
Figure 12: A general NARX model (Hagan et al., 1995b)

The NARX is trained using a single output value, namely the mean squared error over the output nodes. In this study the output nodes contain the water levels at each measuring station. The first output node is the water level at Arnhem, the second at Doesburg, the third at Nijmegen and the fourth at Pannerden. Using the mean squared error over all stations means that minimizing a single station does not help as the best solution. This gives the NARX a single value to fit towards the best options with multiple locations.

The training of these NARX is done using the Levenberg-Marquardt optimization. For this algorithm the FM runs that have been done are divided randomly in a 60,20,20 ratio for training, verification and testing. The training runs are used to change the weights of the NARX. This changing of the weights is done by minimizing the error defined above. This error is the mean squared error over all training runs, timesteps and locations. The verification runs are used to show if including more data improves the NARX, by training the model with both the training and verification data. The testing data is used to validate the NARX in its current training state.

4.3   Discharge only NARX

The most basic approach would be to keep the roughness values constant and only make the NARX on five input parameters, namely four logical arrays if a dike has been breached and the discharge at Lobith. The logical array is an array of the amount of timesteps by the number of breaches containing zeros or ones. The one means the dike is breached, the zero means the dike is intact. To create such a setup, the NARX consists out of five input nodes the ones named above. Further it will have four output nodes, namely the water levels at the different locations; Arnhem, Doesburg, Nijmegen, and Pannerden. The number of hidden nodes is varied to get the best fit in both training and verification. This is done to prevent underfitting and overfitting of the model. The hidden nodes are varied between 1-34 nodes, from which the best result is taken. Between the output and the input there is a feedback loop. This is done by including the result of the previous timestep in the current timestep. Each calculation of the NARX is a given timestep, these are taken at one day.

4.4   Roughness, Discharge NARX

The more advanced approach is, to also include the different roughness values. Therefore, NARX is build using 19 input nodes, namely the 14 roughness sections, four logical arrays for dike breaches, and the discharge at Lobith. The output nodes are still going to be the water levels at Arnhem, Doesburg, Nijmegen, and Pannerden. Also, here the hidden nodes are varied to prevent the over and underfitting of the original model. Next to the hidden nodes the feedback delay is varied between 1 and 10. This means that 3 tests for each combination between hidden nodes and delay is made. The overall functioning remains further identical to the discharge only case.

5   Calibrating discharge

The algorithm for reducing the error between the water levels in 1809 and the surrogated model output, is the barrier function interior point algorithm. This algorithm is capable of optimizing smooth nonlinear problems (Al-khayyal & Sherali, 2000; Byrd et al., 2000). The error in this problem is defined as the root summed squared error. In the function described below the root squared sum error (RSSE), is defined as a function of the value of the surrogated model (y_s) and the measured value in 1809 (y_r).

$$RSSE = \sqrt{\sum(y_s - y_r)^2}$$

The advantage of using the RSSE is that overshoots and undershoots do not cancel each other. This is useful, as the water levels need to match the measured values instead of given a trend. The input values that the algorithm can change differ for the two approaches. For 4.3 this is the discharge wave at Lobith, and for 4.4 this is the discharge wave at Lobith and the roughness sections values. As the algorithm uses a barrier function it can constrain the solutions to several criteria. These criteria are a linear constrain, and a constant constrain. It is known that some of the mathematical solution are impossible in the physical world. The constrains used in the project are constant constrains, namely the manning roughness coefficient should be between 0 and 0.8, and the discharge should be between 6000 and 20000 m^3/s. These values have been chosen as the model is initialised at 6000m^3/s and it is physically impossible to get negative roughness values.

6 Results: Surrogated model accuracy

In this chapter the different surrogated modelling techniques that are used will be discussed, each of these techniques is discussed in its own paragraph.

6.1 NARX accuracy discharge only

According to the description in paragraph 4.3 the NARX model is set up. Using multiple iterations, the correct number of hidden nodes is determined. This is based on the fitting error between the NARX and the DFlow-FM results. Figure 13 shows how the NARX is build up. The **x(t)** includes, the four logical arrays of dike breaches and the discharge wave at Lobith. Node 1-4 give the status of a dike breach at a moment in time, in which 1 means breached and 0 means non-breached. Node 5 gives the discharge at Lobith at the current time step. These inputs and the four output nodes of the previous day get transferred with weights to the hidden layer. The output nodes are the water levels at Arnhem, Doesburg, Nijmegen and Pannerden, found in **y(t)** as node 1 to 4 respectively.
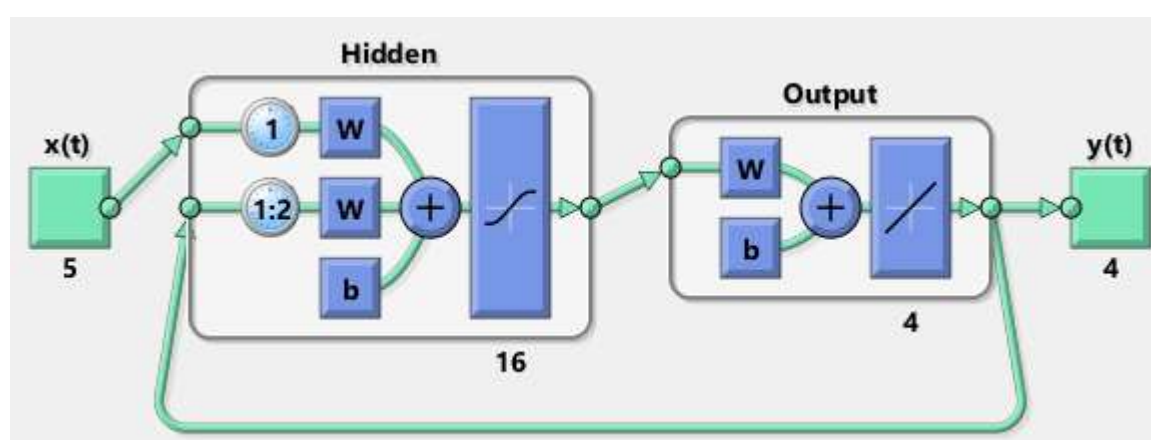
Figure 13: Setup of the NARX for the discharge only

The accuracy of the NARX is determined as the difference between the FM value and the NARX value. This NARX has an accuracy of tens of centimetres at all location except for Doesburg. Table 4 provides an overview of the quality of fitting the NARX to the Dflow-FM data. The best and the worst fits are given in Figure 14 and Figure 15 respectively. These give a visual representation of the correctness of the fit. This means that the NARX is useful, for Arnhem, Nijmegen and Pannerden. As in these locations the R^2 values are above 0.8 and the order of the largest error is within 10's of centimetres. It shows that the fit for Doesburg is a lot worse than for the other measuring stations. A possible explanation for this is the missing roughness values, the Ijssel river has a completely different shape than the others. This will translate in another relation between the discharge and water levels. Whilst three output nodes react similar and one different, it is logical that the fit will be best for the similar reacting nodes. As this will give the lowest error overall.

Table 4: The error between FM and NARX of the discharge only case

Location	Δh best	Δh worst	R^2 best	R^2 worst
Arnhem	4.72 cm	10.15 cm	0.9986	0.8720
Doesburg	6.12 cm	36.50 cm	0.9992	0.0541
Nijmegen	6.09 cm	10.47 cm	0.9973	0.9572
Pannerden	5.32 cm	9.40 cm	0.9988	0.9362
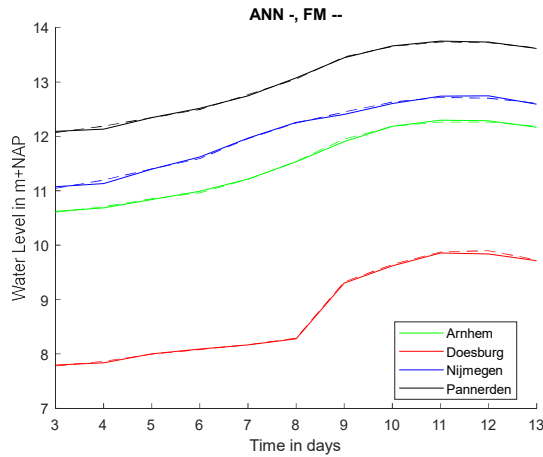
Figure 14: Best Fitted run of the samples
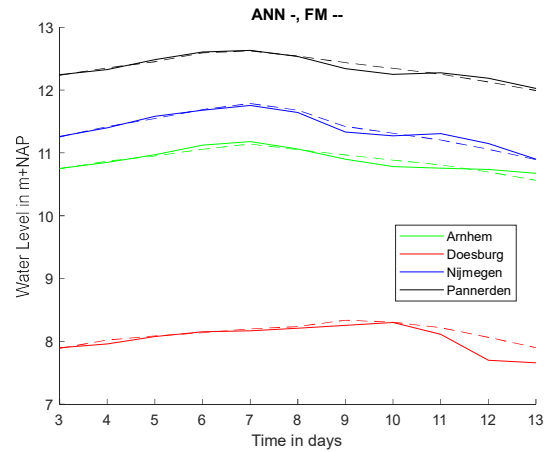
Figure 15: Worst fitted run of the samples

Appendix A has all the 32 Dflow-FM runs. These give an overview of the fitting in all these cases. These show that the behaviour over all the 32 runs is similar to the behaviour seen in Figure 14 and Figure 15. This means that the overall performance is well, only that the performance of Doesburg is the worst off all of them. This also shows that with few variables the surrogated model can already represent the Dflow-FM model.

6.2   NARX accuracy roughness + discharge

According to the description in paragraph 4.4 the NARX model is set up. Using multiple iterations, the correct number of hidden nodes is determined. This is based on the fitting error between the NARX and the DFlow-FM results. Figure 16 shows how the NARX is build up. In the **x(t)** are located, the roughness values, the four logical arrays of dike breaches and the discharge wave at Lobith. Node 1-14 has the representative manning value for the ordered list of roughness sections. Node 15-18 give the status of a dike breach at a moment in time, in which 1 means breached and 0 means non-breached. Node 19 gives the discharge at Lobith at the current time step. These inputs get then put in the hidden layer together with the four output nodes of the previous day. The output nodes are the water levels at Arnhem, Doesburg, Nijmegen and Pannerden, found in **y(t)** as node 1 to 4 respectively.
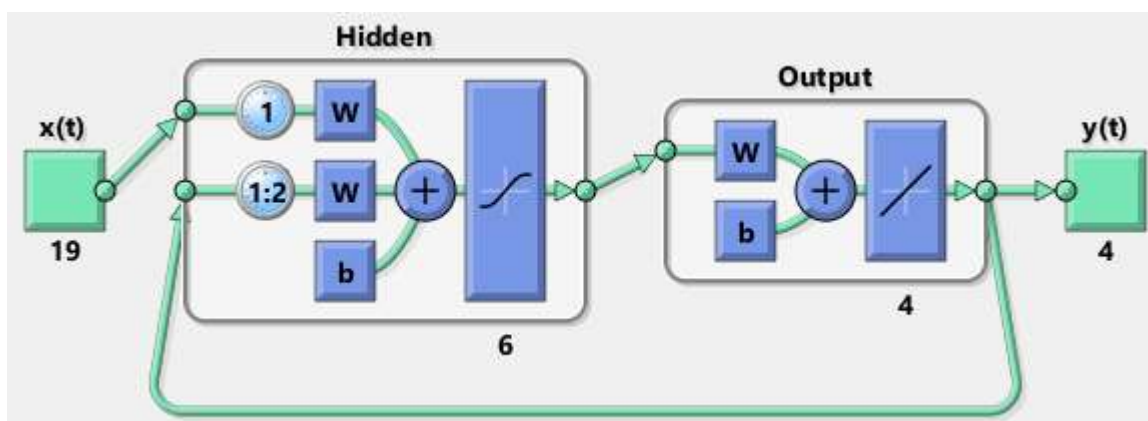
Figure 16: Setup of the NARX roughness+Discharge

Table 5 provides an overview of the quality of fitting the NARX to the D-Flow FM data. The best and the worst fits are in Figure 17 and Figure 18 respectively. These give a visual representation of the correctness of the fit. This means that the NARX is useful, for Arnhem, Doesburg and Pannerden. As in these locations the R^2 values are above or close to 0.8 and the order of the largest error is within 10's of centimetres. It shows that overall the fits are more accurate than in 6.1. This agrees with the different relations between some of the measurement stations and the discharge wave at Lobith. In this it does show that roughness section that effect multiple stations are preferred over single stations. This shows in the difference in accuracy between Arnhem, Pannerden, Doesburg and Nijmegen. As Nijmegen scores worse, but also is physically effected by different roughness sections.
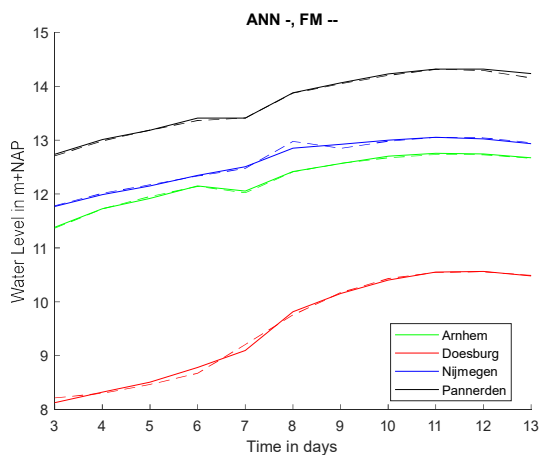
Figure 17: The best run of this set
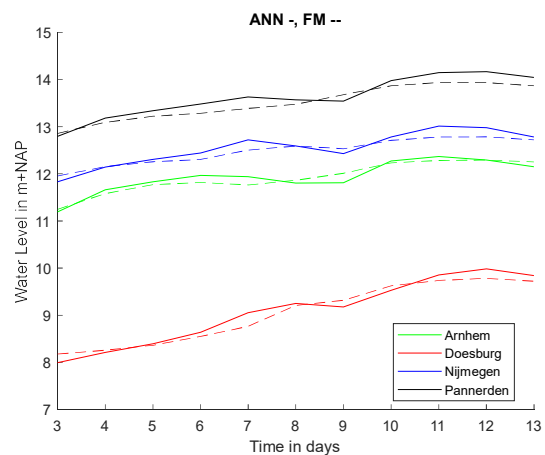
Figure 18: Worst run of this set

Appendix B has all the 32 D-Flow FM runs, this overview shows that all runs perform similar to the best and worse run. As such it is to be assumed that the quality of fit of the system is somewhere between these two runs. Considering that training data will score higher, the worst run gives probably the best overview of applications of the NARX without the training set. This means that for the reconstruction the expected error is probably closer to the worst run than the best.

Table 5: The error between FM and NARX of the roughness and discharge case

Location	Δh best	Δh worst	R² best	R² worst
Arnhem	3.57 cm	19.94 cm	0.9978	0.8852
Doesburg	10.88 cm	29.03 cm	0.9961	0.9438
Nijmegen	12.81 cm	23.35 cm	0.9886	0.7467
Pannerden	4.39 cm	24.22 cm	0.9963	0.7888

7 Results: Discharge during 1809 flood event

7.1 Changing only the discharge

Using the trained NARX from 6.1 and the method of paragraph 5, the discharge wave at Lobith is changed to match the measured water levels. The discharge wave can change by independently changing the discharge at each day. Figure 19 and Figure 20 provide the results of this optimization sum. This result is created by using 32 different initial conditions. The initial condition for each case is one of the 32 runs out of FM. It shows that the missing persistence in the discharge wave allows for unphysical changes in the discharge. This might be solved by only allowing known discharge waves to scale or move. The matching to the water levels stays far off. This can mean that the initial roughness values are wrong. This is then compensated by a higher discharge. Due to the delays in the NARX it shows that the peaks of the discharge and water levels do not meet correctly. However, in the physical domain there is travel time between this wave and the measurement locations. Furthermore it shows that the overall error is minimized by having most of them wrong but in the correct trend.
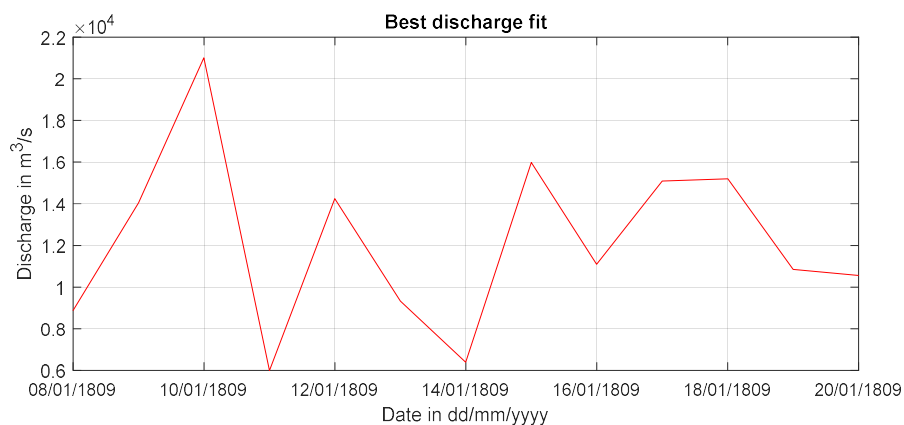
Figure 19: Best fitted discharge
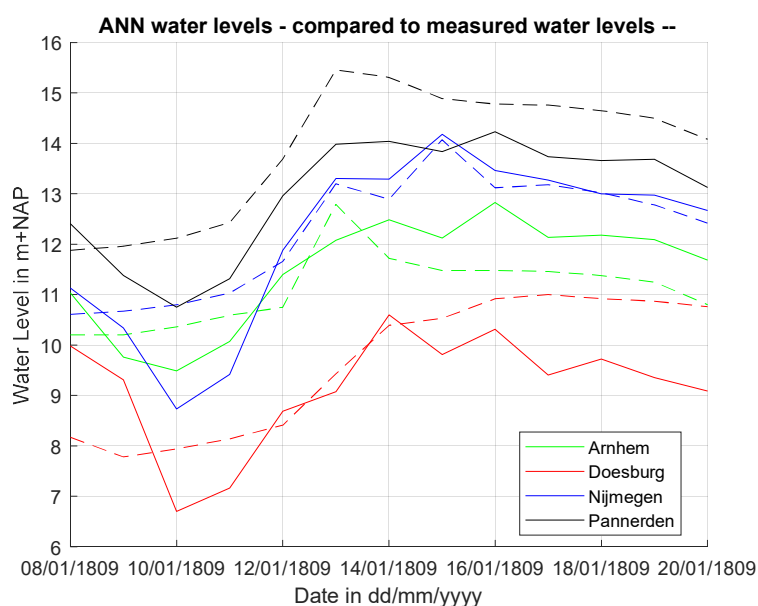
Figure 20: Water levels with this discharge

7.2 Roughness + discharge

Using the trained NARX of 6.2 and the method described in paragraph 5. The discharge wave at Lobith and roughness are changed to match the measured water levels. Figure 21, Figure 22, and Figure 23 provide the results of this optimization sum. This result is created by the barrier function interior point algorithm with the error being the RSSE of the designated water levels in this case Arnhem and Doesburg. The discharge wave shows a possible form from day 3 onwards. This seems promising however can be a random lucky result. Considering the water levels at Arnhem and Doesburg the errors are quite large. Especially for the station of Arnhem the trend of the graph is different, and the water level has errors over 2m. This means that the result cannot be considered accurate for the given case. Again, this shows that the roughness values in the FM model are probably completely off and this is being compensated by the discharge.
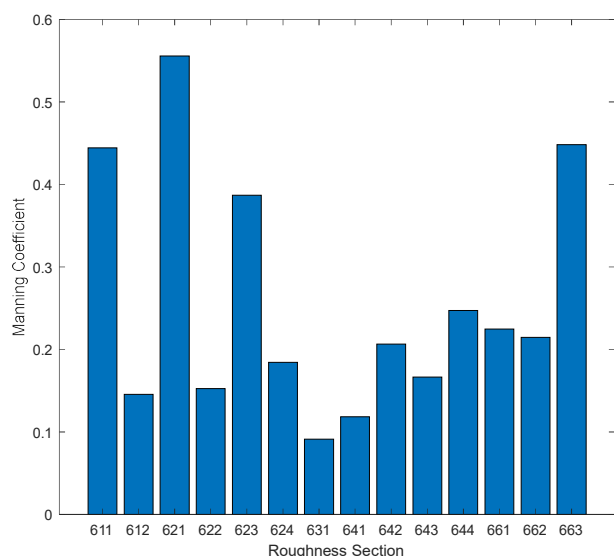
Figure 21: Optimized roughness for the 1809 case
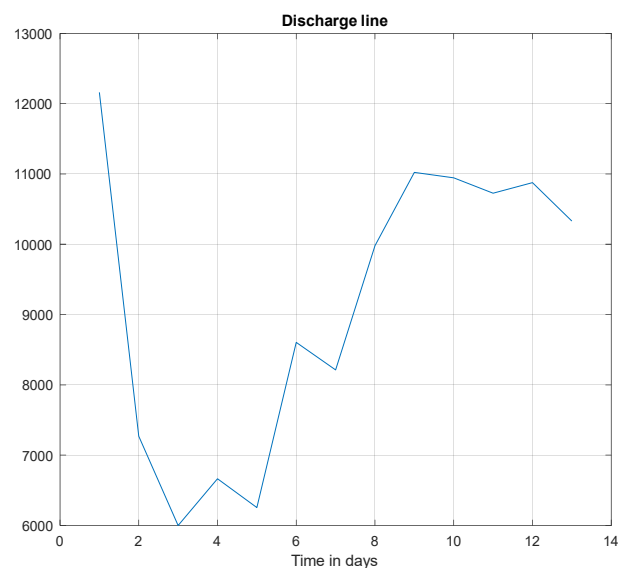
Figure 22: Discharge wave of the 1809 case
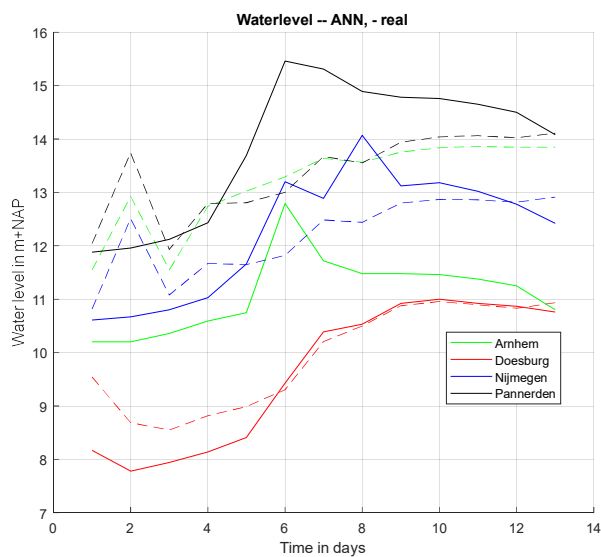
Figure 23: Water levels after optimizations

7.3 Discharge wave by only moving and scaling

In the past two actions it showed that the persistence of the discharge wave is lost. Therefore, a scaling approach is tried for the discharge wave. This scaling procedure on all the measuring stations, shows a peak discharge of 60000 m³/s (Figure 24). Figure 25 shows the corresponding roughness values for this discharge wave, to get the water levels of Figure 26. These water levels show that the fit is poorly. Even though being the best fit the algorithm could find, the water levels in the peaks are most of. The run up and backside of the curve are accurately represented, especially for Arnhem and Nijmegen. Looking at Pannerden the overall picture is a lot worse, with errors of more than 1m. This shows that the roughness values or other criteria not included should influence this relation. As these are then incorrect, the method shows its limitations. The location Doesburg shows limitation in the first days, only around the peak the values are correct.
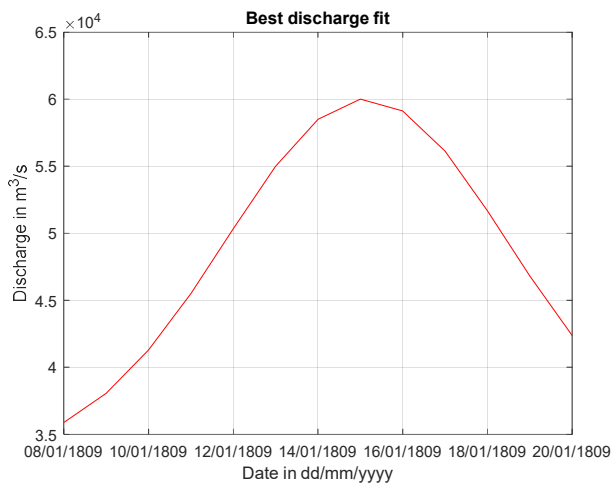
Figure 24: Discharge wave after optimization
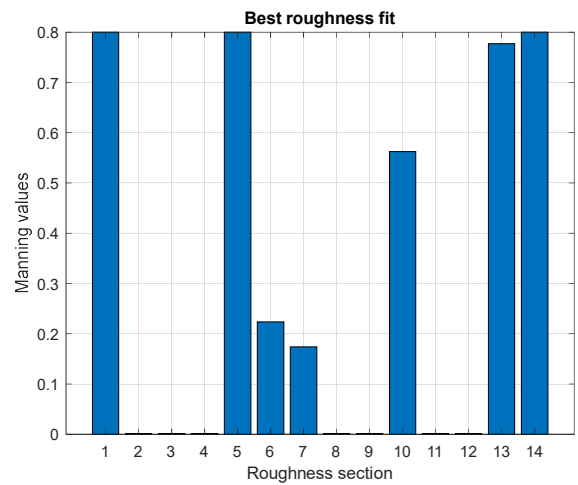
Figure 25: Roughness values after optimization
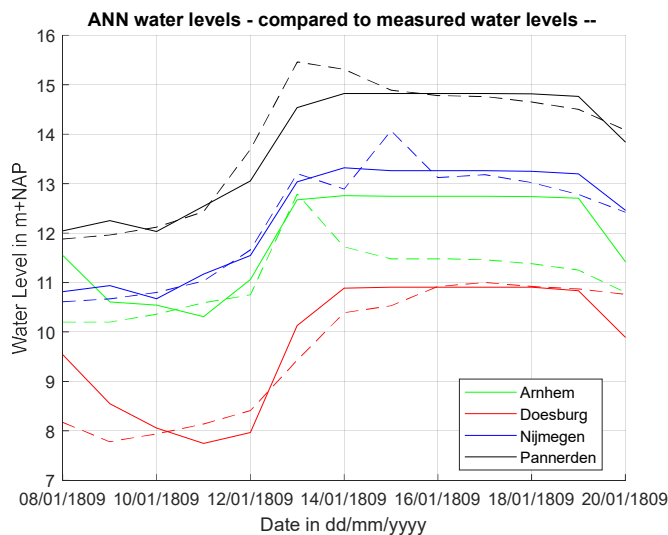
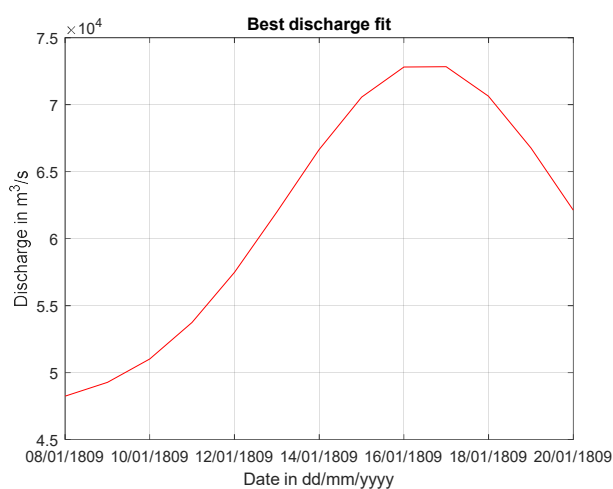Figure 26: Water levels after optimization

28

8 Sensitivity analysis of uncertain parameters

To ensure the robustness of the methodology, a sensitivity test has been done on several factors. These factors are the exclusions of measuring stations and manipulation of the measured data. The manipulation of the measured data is done to allow for changing measuring times. This is done by increasing or decreasing the measured daily level till 10% change. Some measured stations are excluded as these should not affect the eventual conclusions.

8.1 Using different amounts of measuring stations

Using the different measuring stations to reconstruct the discharge, it should yield the same discharge function. This hypothesis is tested using different scenarios of inclusion and exclusion of stations in the error definition of the optimization function. This error is defined as the RSSE, in which the sum of all stations is considered. With the exclusion of a station, this station is no longer part of the summed value.

8.1.1   Doesburg

Using only one measuring station, the one at Doesburg, it shows that the optimization function can reconstruct a discharge wave. Figure 27 gives the discharge wave after the optimization function has found an optimum, this discharge wave however is way larger than anything that the literature has found till now (Toonen, 2013). The optimal solution is shown in Figure 29 this shows that the fit to the measured data is accurate. It shows some small errors in the run-up phase, but the peak and end timing is accurately represented. However, it does show that the other stations, not included in calculation, score significantly worse. First, is the timing off by several days, and is the shape different in comparison to the measured values. These can probably be explained by incorrect roughness values, shown in Figure 28, whilst these stations were not included in the error calculation. Second, it shows that the sharp peaks in the other stations are not represented. These could occur due to the ice dams that were present during the event but missing in the trainings data and as such in both the surrogated and the physical model. Due to these missing values the extreme peaks are missing in the surrogated output, and the model cannot reproduce this input. Third, the water level errors in the run up can be caused by having insufficient data points in time, this leads to an underfitting of the model.
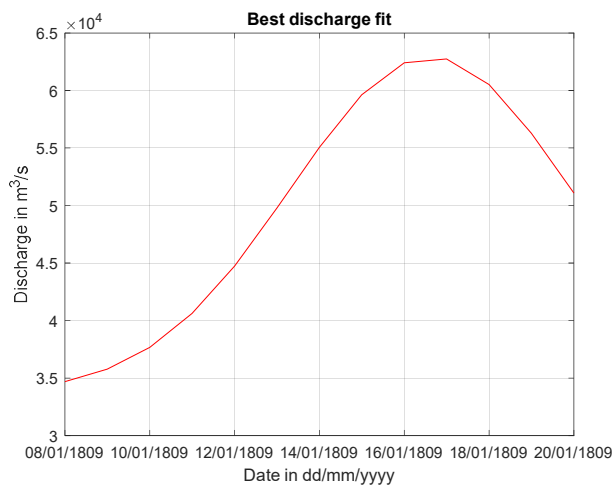
Figure 27: Discharge wave at Lobith
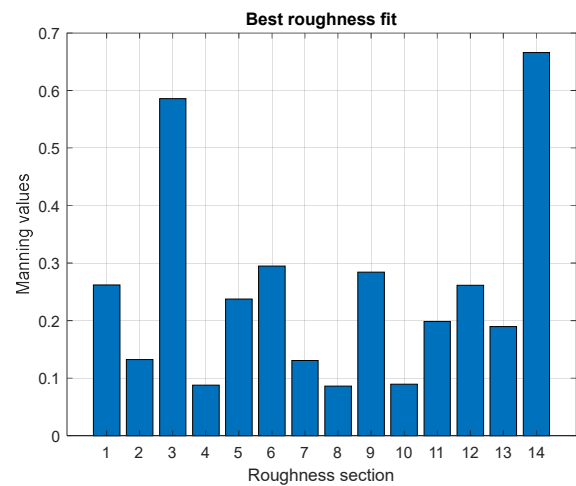
Figure 28: Roughness sections values
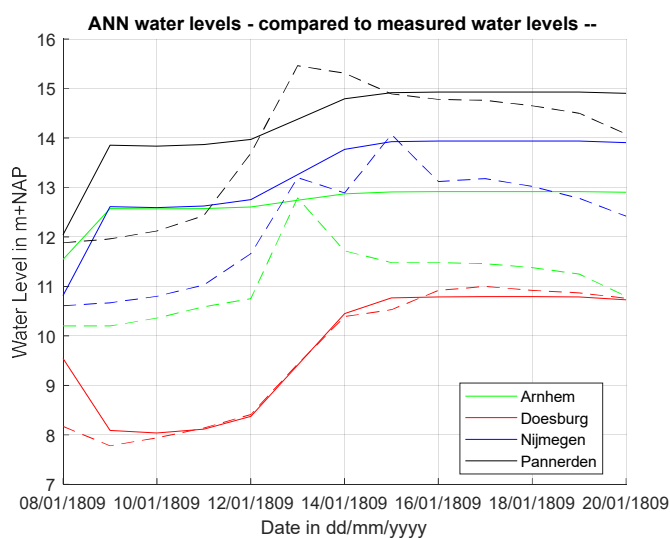
Figure 29: Water levels after optimization

8.1.2 Doesburg & Pannerden

Using the two measurement stations spatially separated, namely Doesburg and pannerden. The discharge wave is reconstructed. Figure 32 shows the results of the best fitted discharge wave (Figure 30), with corresponding roughness values Figure 31. It clearly shows still a large error in the measured water levels and the NARX levels, at all measurement location. This means both the included and excluded stations. Doesburg is the closest approximation in this result. Considering the discharge wave of Figure 30, this shows unrealistic values for the discharge at Lobith. Also considering the roughness values of Figure 31 is not realistic. All the values are outside of the theoretical range. This is not strange for a model but makes the result more unbelievable. Considering these points, the number of data points are too few to find the correct results, with the given algorithm.
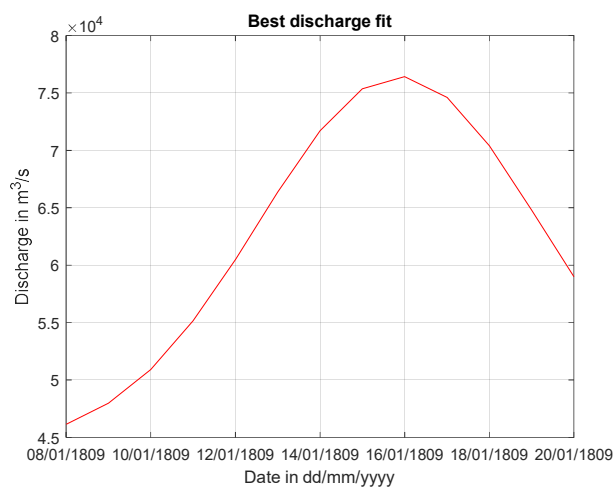
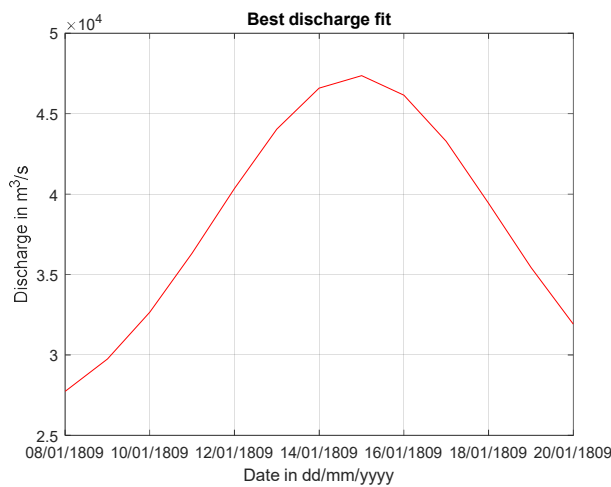*Figure 30: Discharge wave after optimization*
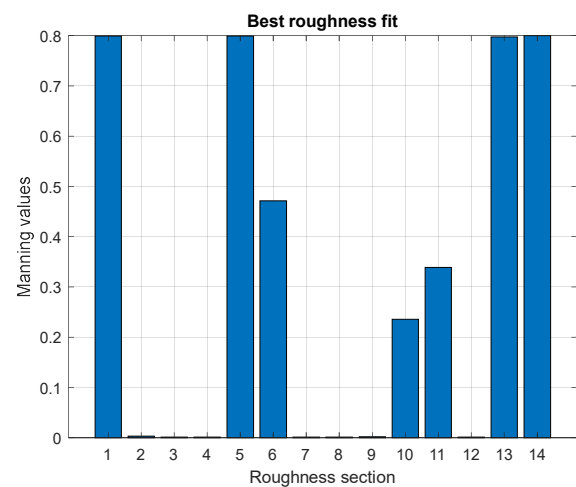
Figure 31: Roughness values after optimization
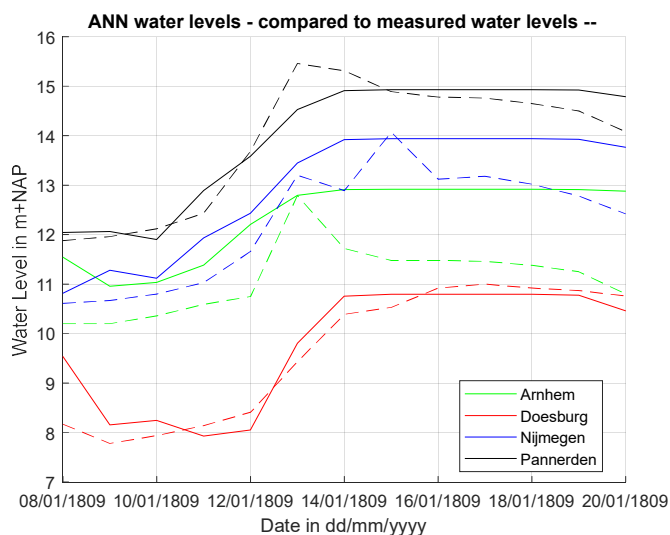
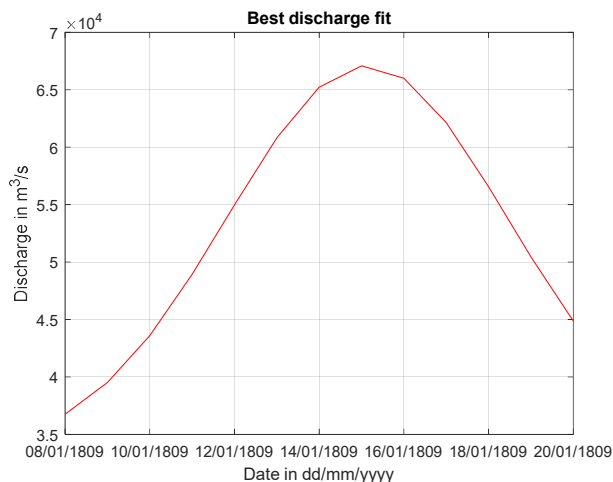*Figure 32: Water levels after optimization*

8.1.3   Arnhem & Nijmegen

Using two measurement stations, the same results should be achieved. This time the stations Arnhem and Nijmegen are considered. It shows that the result is quite accurate for Nijmegen, however not at all for Arnhem, see Figure 35. From this it shows that the optimization algorithm is not able to reach a correct discharge wave. This can be caused by having to many variables in comparison to the data or using an algorithm that is not correct for the given problem. Figure 34 shows extreme roughness values, something that strengthens the believe that the optimization algorithm cannot find the correct values. Figure 33 shows the discharge wave resulting from the optimization this is more extreme than any value in the literature. As such the method seems to show randomness in its working.
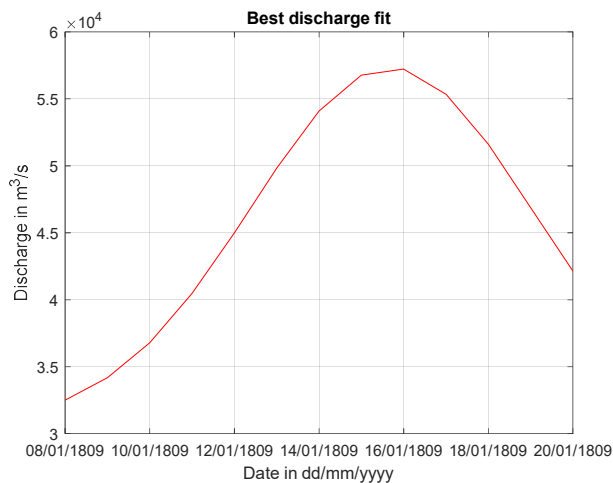
Figure 33: Discharge wave after optimization
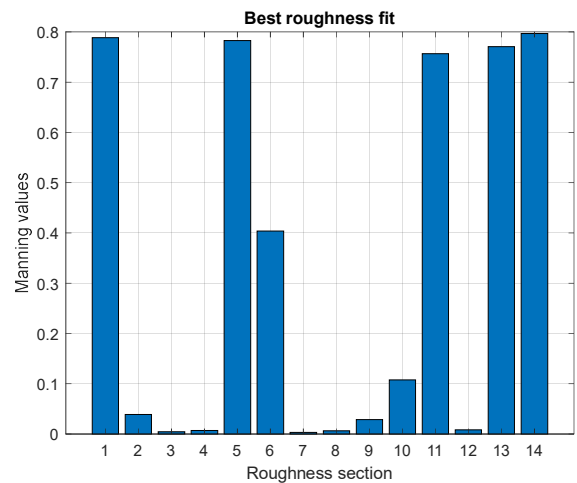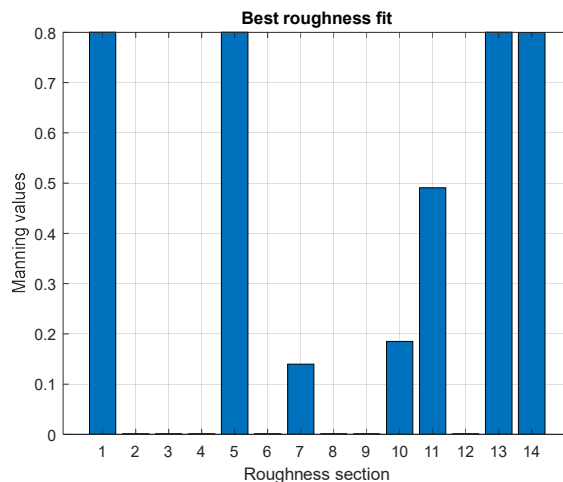

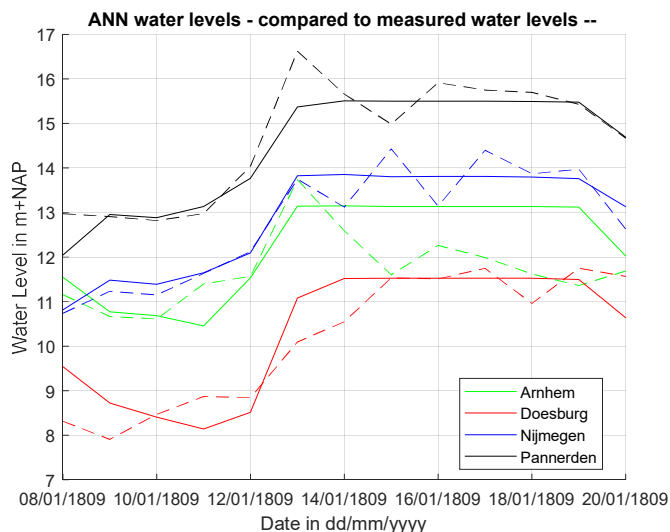*Figure 34: Roughness values after optimization*
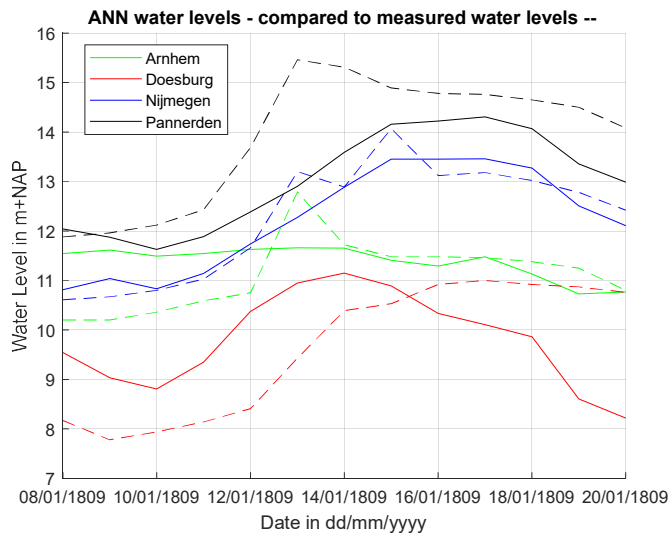
Figure 35: Water levels after the optimization

8.2 Changing the real water level random till +10%

With data of 1809, the timing of data is uncertain. As such the data is manipulated in this chapter to be increased for each day between 0 and 10%. The increased values are to allow for measurements taken earlier than expected in the run up and measurements taken later than expected in the run off.

8.2.1 Doesburg

Figure 36 shows the discharge wave after optimization with the error of only measuring station Doesburg. This discharge wave has a peak discharge of 70,000 m^3/s, which is way larger than any value mentioned in the literature. Using Figure 38 it shows that the fit of the water levels is nice and accurate. However, this is only the case for Doesburg, all the other stations have a completely different reaction. The roughness values for this case are shown in Figure 37, these are also strange. This can be caused by having to few data points in comparison to the number of variables.
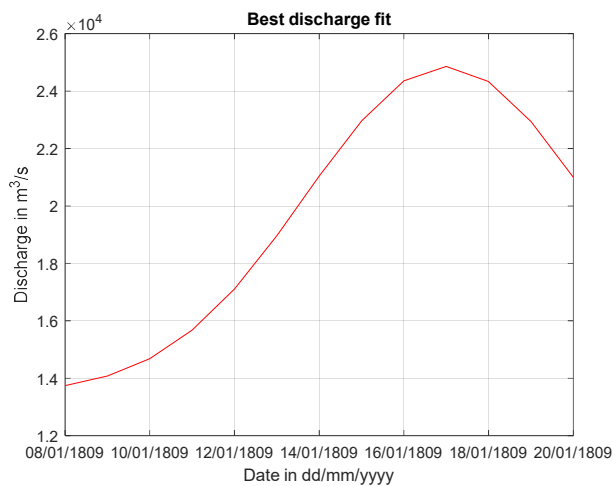

Figure 36: Discharge wave after optimization
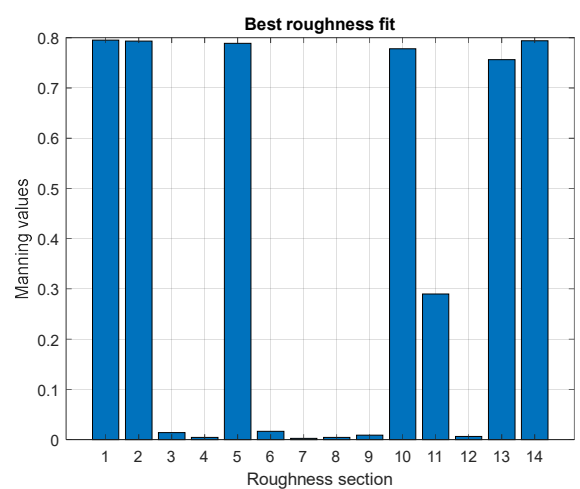
Figure 37: Roughness values after optimization
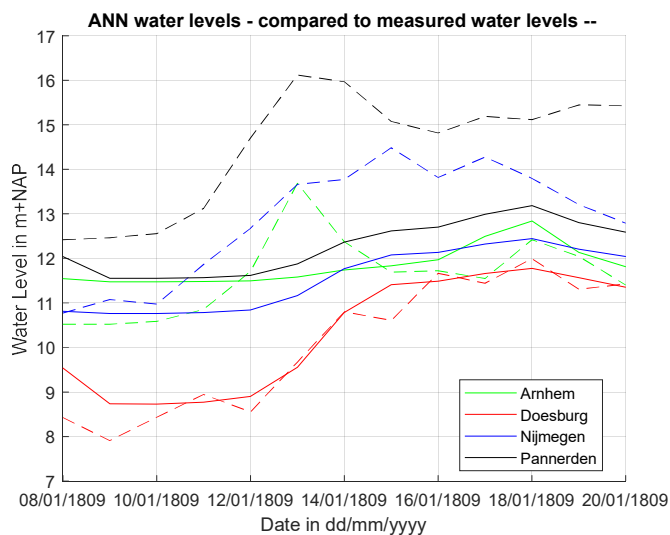
Figure 38: Water levels after optimization

8.2.2   Doesburg & Pannerden

Figure 39 shows the discharge wave after optimization on two stations namely Doesburg and Pannerden. One of the interesting effects in this case is the topped of water levels found in Figure 41, the causes for these maximum effects are unclear. One potential cause of this is the difference roughness values together with the logical array of dike breaches. The roughness values after the optimization are shown in Figure 40. This shows that the model does find a maximum value, however the model needs this maximum now to reach the measured levels. This shows that there is an error in this method again occurring trough either to few data points or being outside of the surrogated models reach.
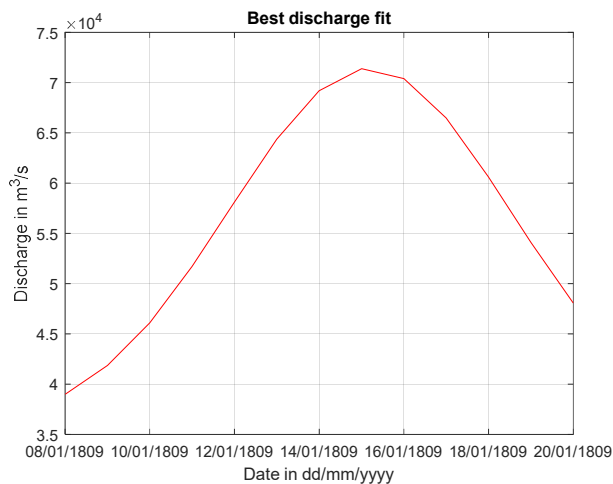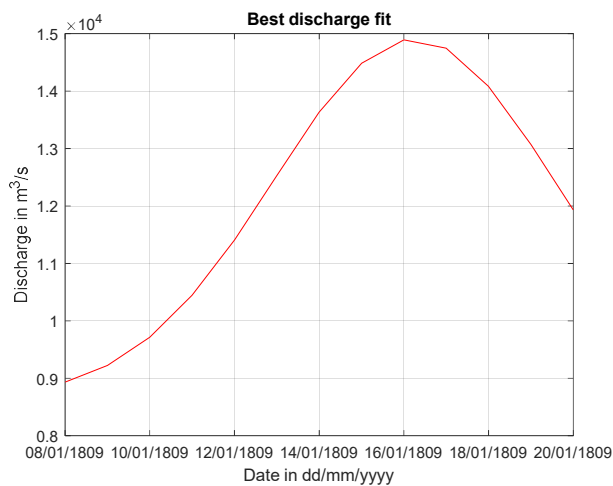

Figure 39: Discharge wave after optimization
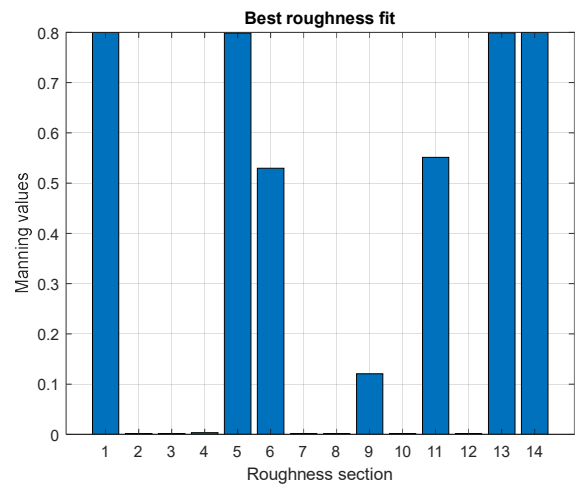
Figure 40: Roughness values after optimization
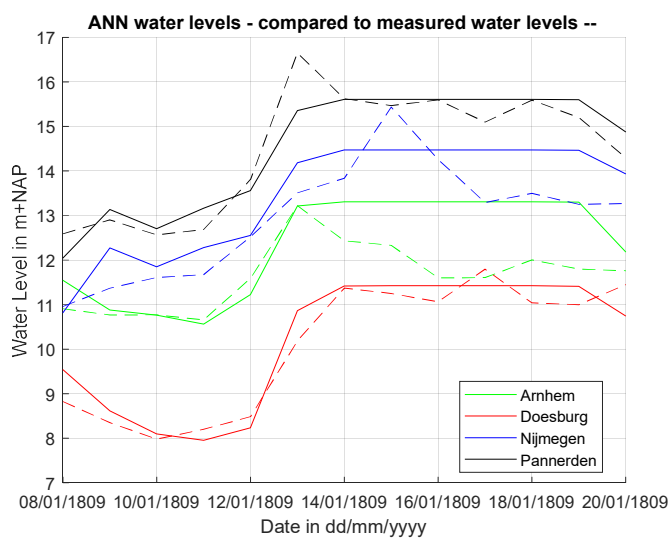
Figure 41: Water levels after optimization

8.2.3 Arnhem & Nijmegen

Figure 42 contains the discharge wave for the situation measured at Arnhem and Nijmegen. The resulting water levels are shown in Figure 44. These show that the fit is poorly in both Arnhem and Nijmegen, as well as the other measuring stations. It has the interesting result that Arnhem has a correct fitted run up section and Nijmegen has a correct ending of the discharge wave. This means that the overall error seems to be low with the one focussing on the start and the other focussing on the end. Figure 43 shows the roughness values for the given optimal fit. These roughness values are all outside the training range of the surrogated model and together with the extreme large discharge wave, are these results inaccurate and unreal. This shows that either the number of unknowns is to large or the optimization algorithm is unable to find a correct global minimum for the given problem. This can be caused by the surrogated model being outside of the training sphere or the amount of data is insufficient to meet both the number of unknowns and the constrains.
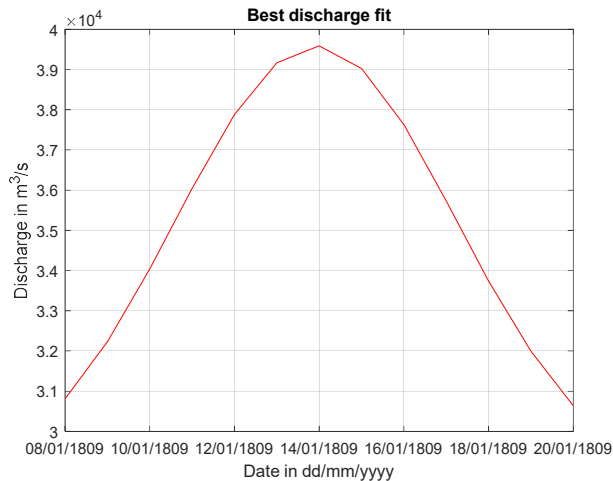

Figure 42: Discharge wave after optimization
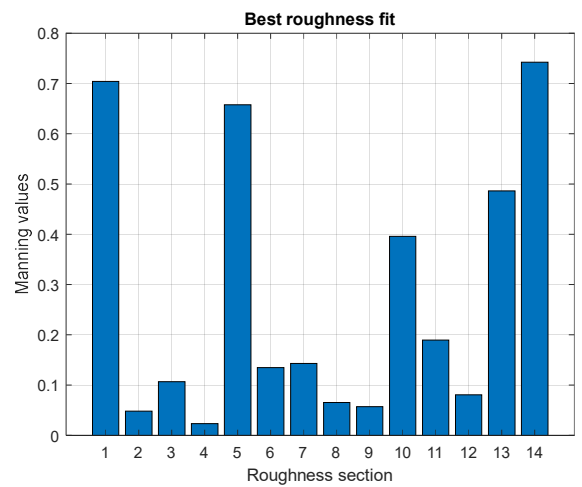
Figure 43: Roughness values after optimization
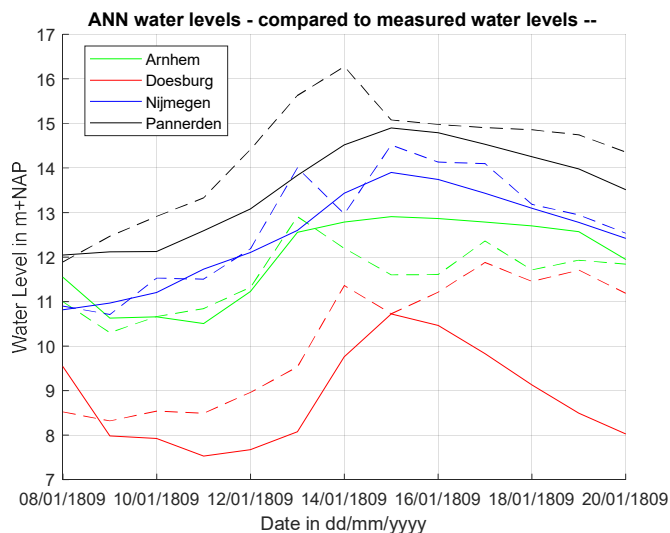
Figure 44: Water levels after optimization

8.2.4 All stations

Figure 45 shows the discharge wave that fits best by the water level of all stations in this scenario. Together with the roughness values displayed in Figure 46. This leads to the water levels described in Figure 47. These water levels show a maximum water level independent of the discharge wave as it is at max for multiple days. This can be caused by the breaking dikes, combined with the high discharge values and the weird roughness values. All in all, does it show that the model is not able to reconstruct this situation. This can have multiple causes as described earlier, in missing data or a optimization algorithm with the wrong constrains, methodology or other issues with the surrogated model not representing the correct solution.
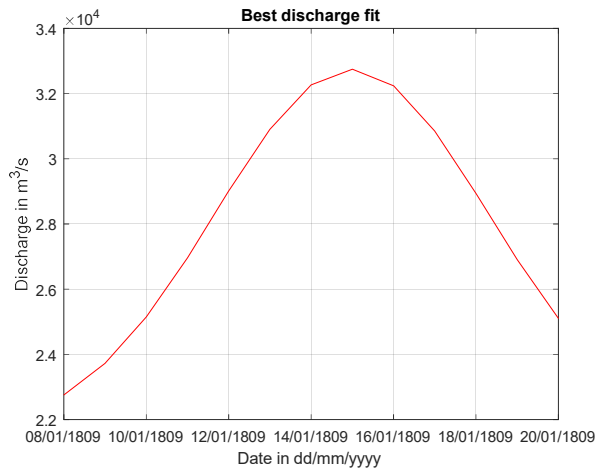

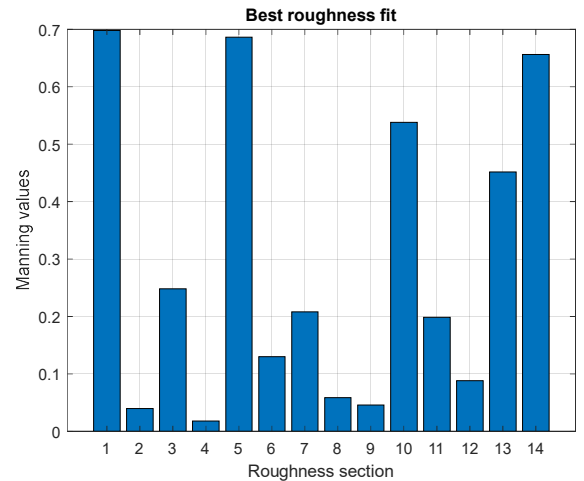*Figure 45: Discharge wave after optimization*

Figure 46: Roughness values after optimization

Figure 47: Water levels after optimization

8.3 Changing the real water level random till -10%

With data of 1809, the timing of data is uncertain. As such the data is manipulated in this chapter to be decreased for each day between 0 and 10%. The decreased values are to allow for measurements taken later than expected in the run up and measurements taken earlier than expected in the run off.

8.3.1 Doesburg

Figure 48 shows the discharge wave that was the input of the best fitted water levels (Figure 50), the other input for this sum are the roughness values shown in Figure 49. The water level fit shows an overall well-defined trend; however, the errors are large at all time steps. Therefore, the concluding inputs are incorrect, in addition to the 1809 flood. This time only Doesburg's error was considered, which is clearly visible in the results as all other stations are not even close to the measured values. The causes of these errors are partly explainable by a shortage of independent data. The error is calculated for each day, however the change of today will affect tomorrow and the other way around. This persistence in the known measurements leads to false optimums that meet only sections of the wave.
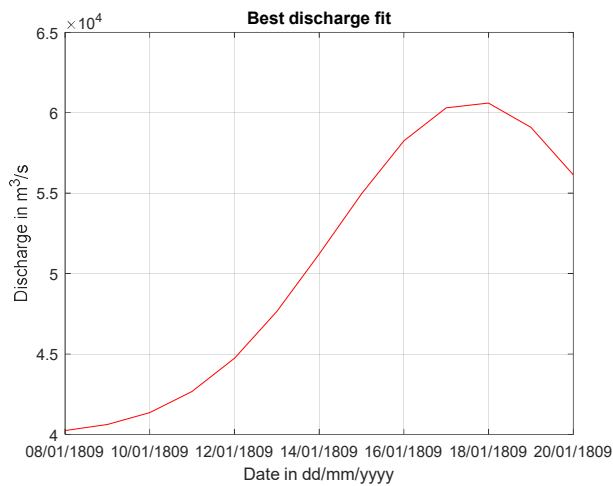

Figure 48: Discharge wave after optimization
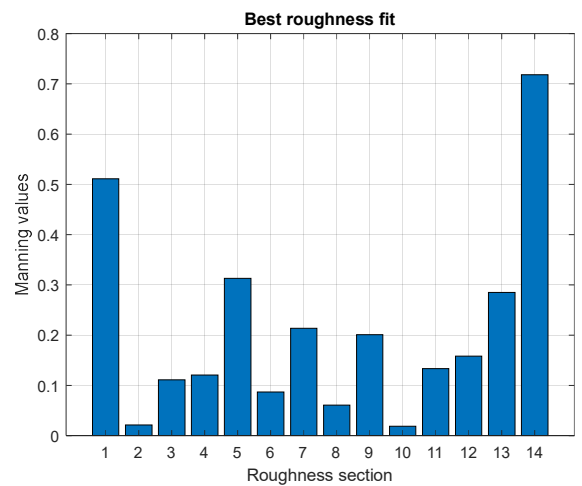
Figure 49: Roughness values after optimization
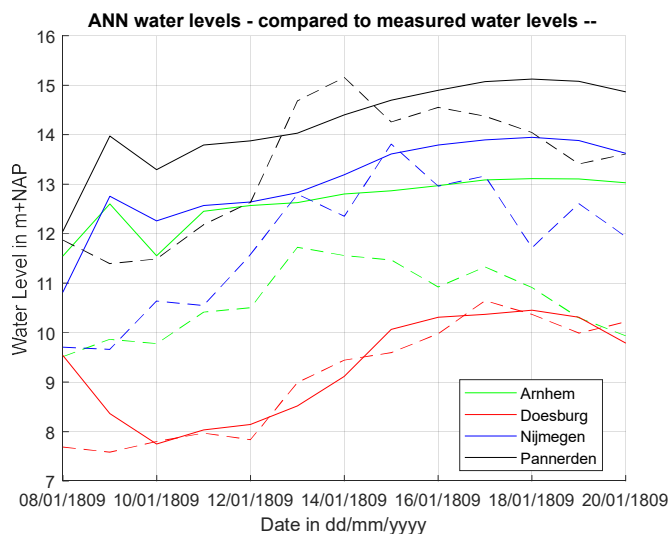
Figure 50: Water levels after optimization

8.3.2 Doesburg & Pannerden

The inputs discharge wave and roughness values for the best fitted version for Doesburg and Pannerden are shown in Figure 51 and Figure 52 respectively. The resulting water level are shown in Figure 53. This shows that Pannerden overall scores nicely, in following of the measurement and having a small error, except for the peak around the 13th of January. However, Doesburg has a large error, which is incredible larger than Pannerden and Nijmegen. This is weird as Doesburg was part of the error calculation whereas Nijmegen is not. This can be caused by the small difference in effects of Nijmegen and Pannerden.

Figure 51: Discharge wave after optimization
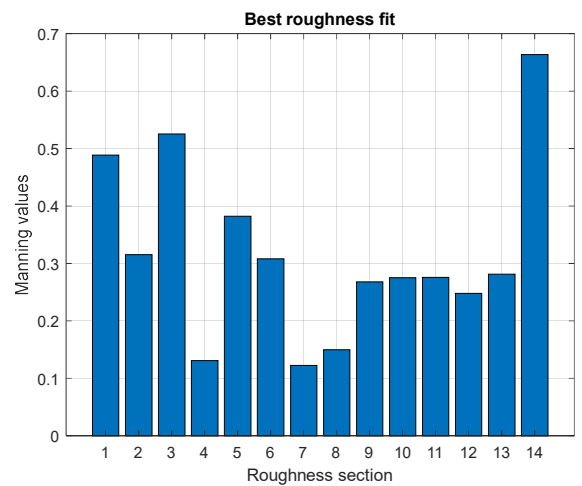
Figure 52: Roughness values after optimization
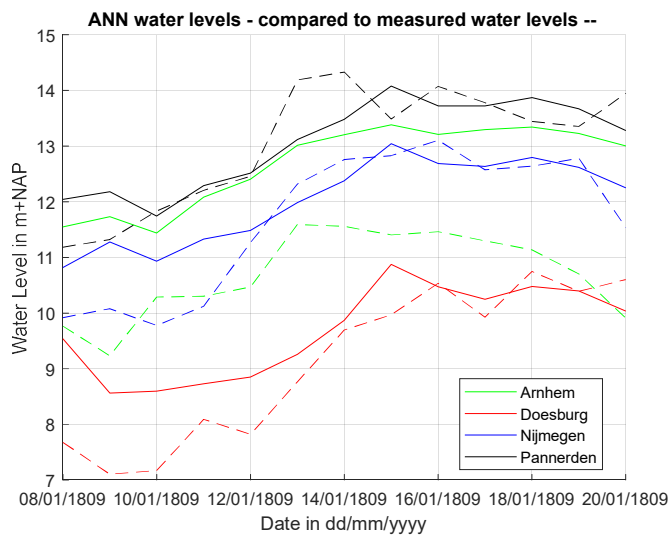
Figure 53: Water levels after optimization

8.3.3 Arnhem & Nijmegen

Figure 56 shows the resulting water levels with the discharge and roughness defined in Figure 54 and Figure 55 respectively. This shows mostly a bad fit of both Nijmegen and Arnhem, which has about an equal error in positive values as negative values. The accuracy of this fit is poor. Both Nijmegen and Arnhem fit poorly to the measured values. The errors are over 2m at some places, and around 1m in average. These errors are to large to consider this result correct for the 1809 flood. This can have several causes, being not the best fitted surrogated model for these locations. Next to this the potential of ice dams in the Waal and Nederrijn which are not part of this FM model, but are present in the measurements.
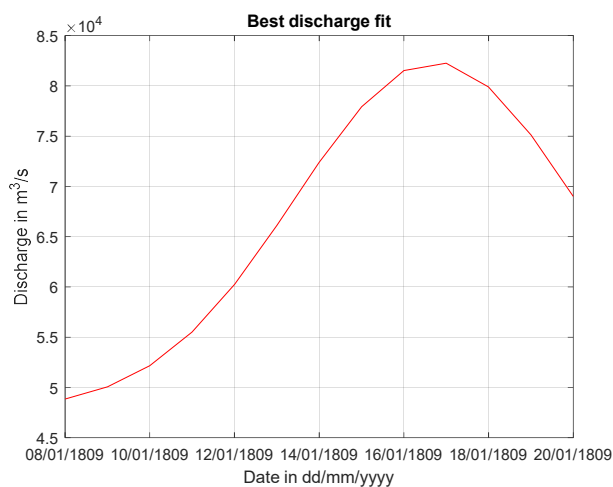

Figure 54: Discharge wave after optimization
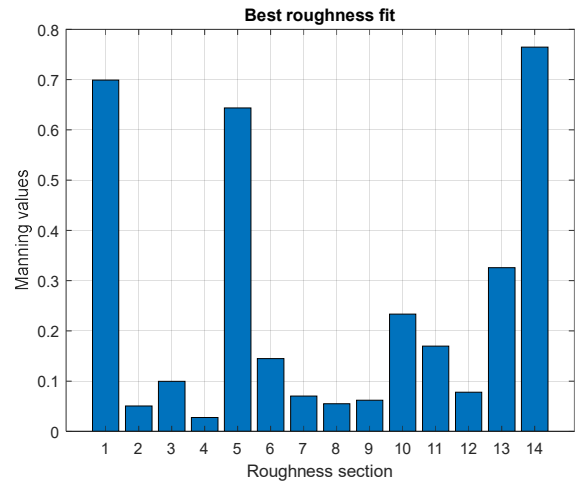
Figure 55: Roughness values after optimization

.
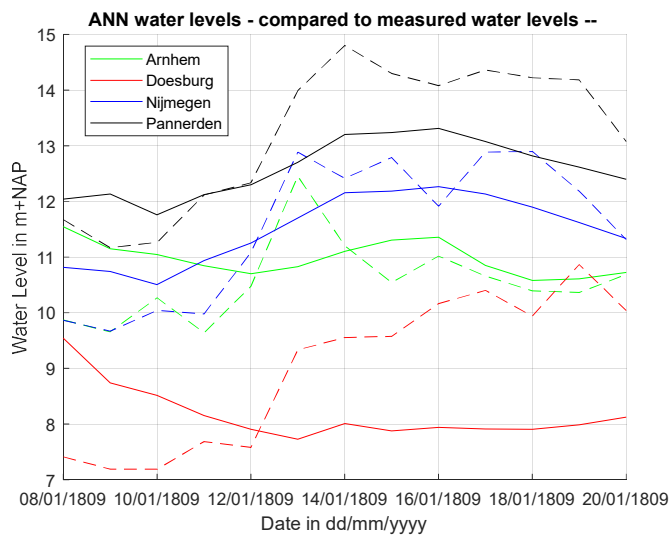
Figure 56: Water levels after optimization

8.3.4 All stations

Figure 57 and Figure 58 are the input variables that lead to the best fit of the water levels in Figure 59 for all stations. The discharge wave and the roughness values are shown in Figure 57 and Figure 58 respectively. It shows in the water levels that the fits overall are in line with the measured values. However, there are large differences from day to day, with over and underestimations. These are most obvious in Pannerden, Nijmegen and Doesburg. In Arnhem this is even not the case, so this leads to a bad fit overall. Mostly the direct changes in the measured water values are not resulting in a change in surrogated models. It becomes clear that these direct changes are unable to be represented in the surrogated models.
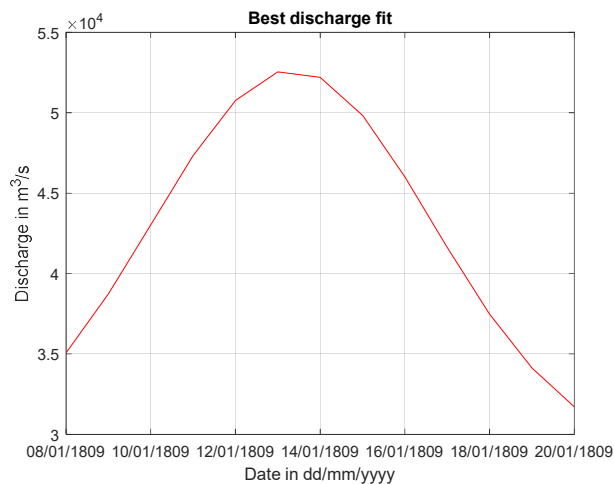

Figure 57: Discharge wave after optimization

Figure 58: Roughness values after optimization
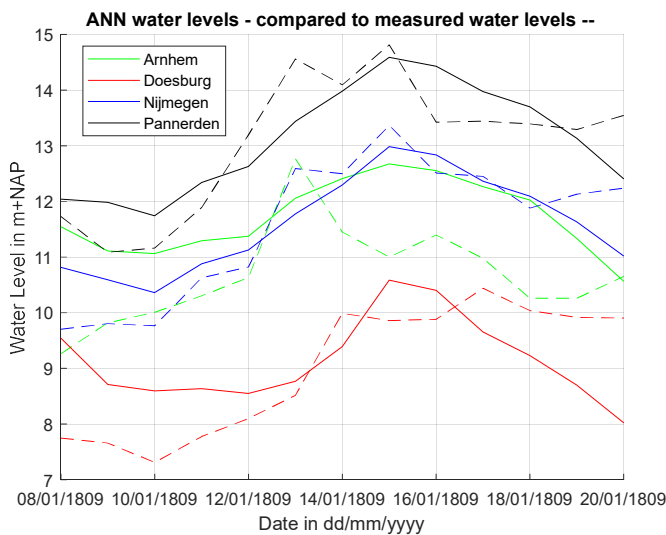
Figure 59: Water levels after optimization

8.4 Changing the real water level random -10% to 10%

The exact times of measurement is unknown of the measured values in 1809. To include changing measurement times the values of the measurements are randomly increased or decreased by 10%. This relates to earlier or later measurement then expected. This situation is tested for different measurement stations included in the calculation.

8.4.1 Doesburg

Figure 60 and Figure 61 are the discharge and the roughness values respectively which lead to the water levels shown in Figure 62. As the water levels show, the fit of the water levels at Doesburg is poor. It has a maximum error of about 1m in the water levels, when ignoring the initial days. All other stations are worse, but this is to be expected as they are ignored in determining the discharge and roughness values. This shows that the method does not seem to work considering the error is given in these cases. These are probably caused by the algorithm not finding the correct minimum situation.

Figure 60: Discharge wave after optimization
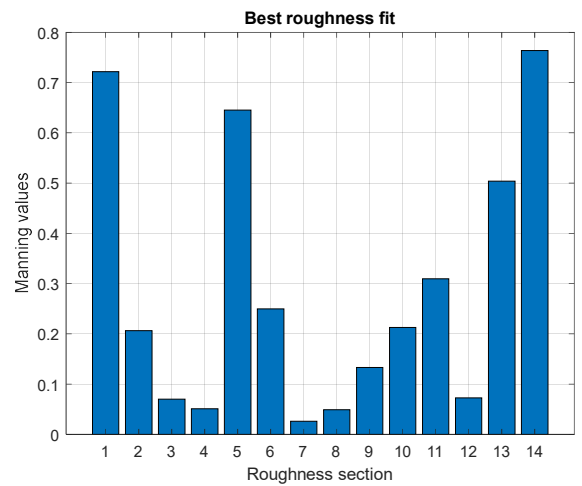
Figure 61: Roughness values after optimization
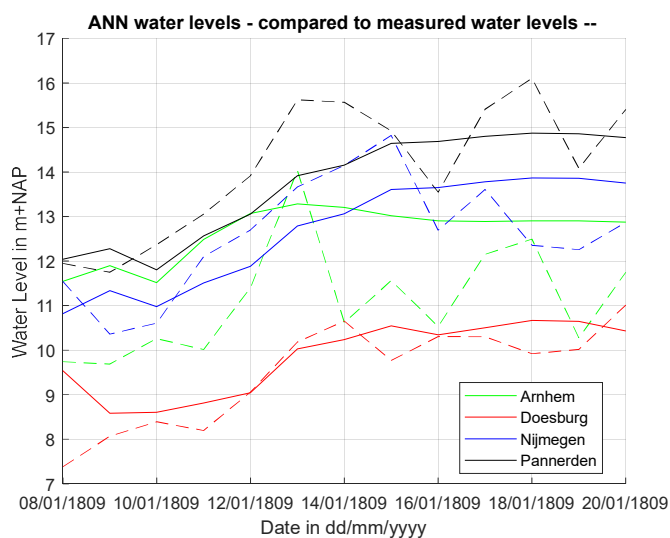
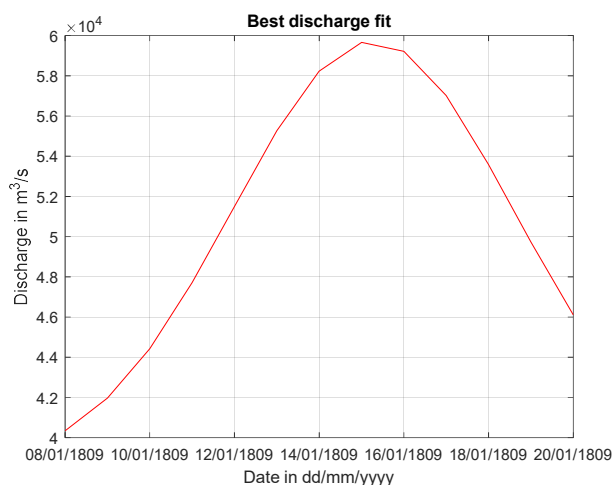*Figure 62: Water levels after optimization*

8.4.2 Doesburg & Pannerden

The water levels presented in Figure 65 are determined by changing the discharge wave and the roughness values which are presented in Figure 63 and Figure 64 respectively. It shows that the general curve of the water levels is followed by these input variables. However, all out of line lying water level values are not representable by the surrogated model. The roughness values as described are mostly reaching the maximum value set in the algorithm. Next to this part is it weird that the roughness sections have large different values in follow up sections. The discharge wave shows extremely high values more than which have been written in any literature. The water levels are for all other stations completely out of line, with errors that are larger than 2m. These levels are expected to be wrong while not being included in the error calculations for the recreation of discharge waves.

Figure 63: Discharge wave after optimization
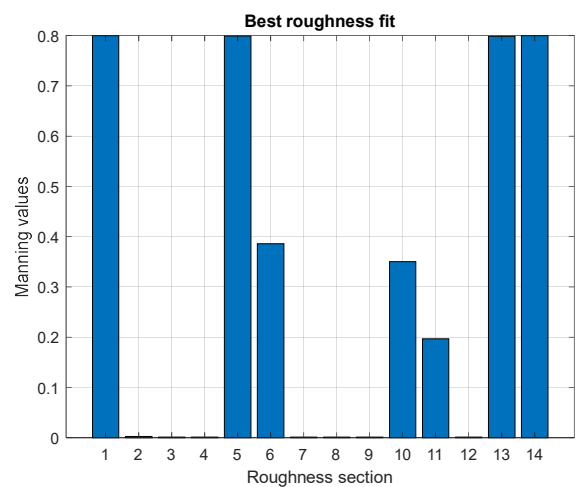
Figure 64: Roughness values after optimization
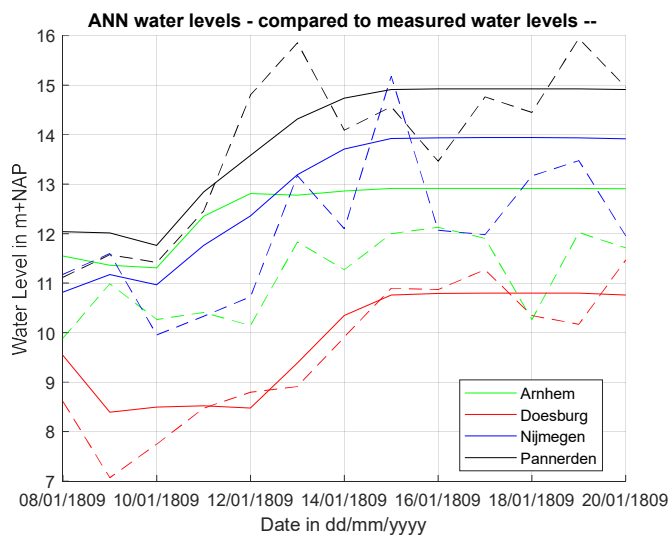
Figure 65: Water levels after optimization

8.4.3   Arnhem & Nijmegen

Figure 66 and Figure 67 represent the discharge wave and roughness values respectively, these are used to reconstruct the water levels shown in Figure 68. De water levels at both Arnhem and Nijmegen, are overall minimized in error. However, there are a lot of overshoots and undershoot between the fits and measured data. The roughness values, that have been used for the water levels show weird values as the difference between the sections are large. Next to this is the discharge wave extremely high above any flood that has been described in the literature. Again the system probably does not functions, due to either being outside the surrogated model range or the algorithm has to few data.

Figure 66: Discharge wave after optimization
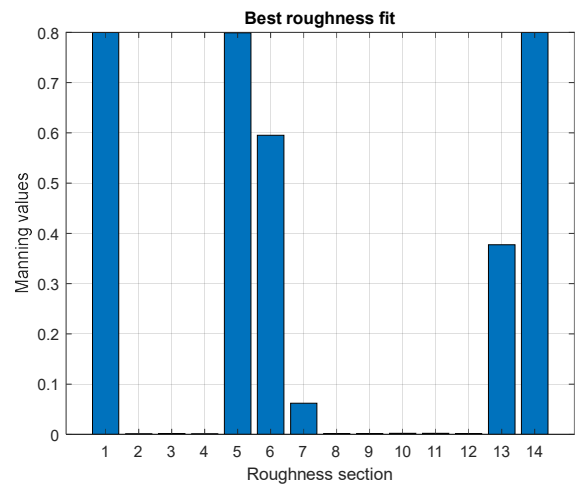
Figure 67: Roughness values after optimization
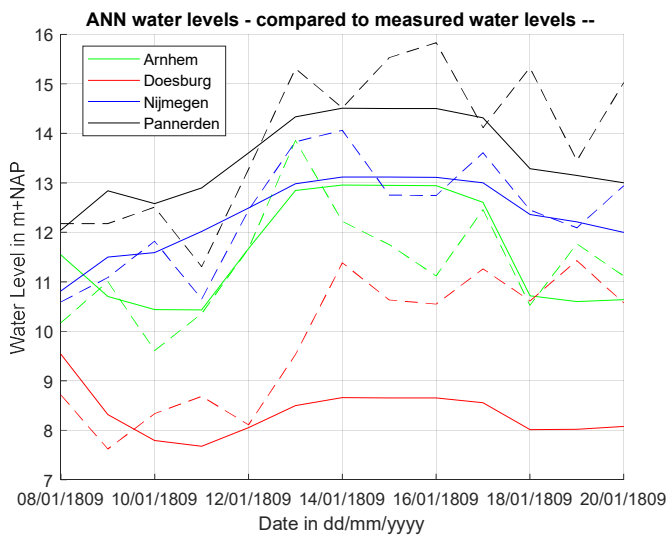
Figure 68: Water levels after optimization

8.4.4 All stations

Figure 71 shows the water levels that are the best fit to the measured values. This is fitted to all the measured stations. It shows that there is a maximum water level that can be reached. This shows in the flatted top. The fit of these water levels is poor, especially in Pannerden, where errors are present of more than 2m. The water levels are generated using the discharge wave shown in Figure 69 and the roughness sectional values shown in Figure 70. These roughness values are extreme in the large changes between different section, namely being either the maximum or minimum value. The discharge wave is way larger than any wave presented in any literature. Furthermore does it show a topping off effect in the water levels that is not present in this discharge wave, which is interesting considering that physically this means that water remains in the system, or gets entered via different routes.

Figure 69: Discharge wave after optimization
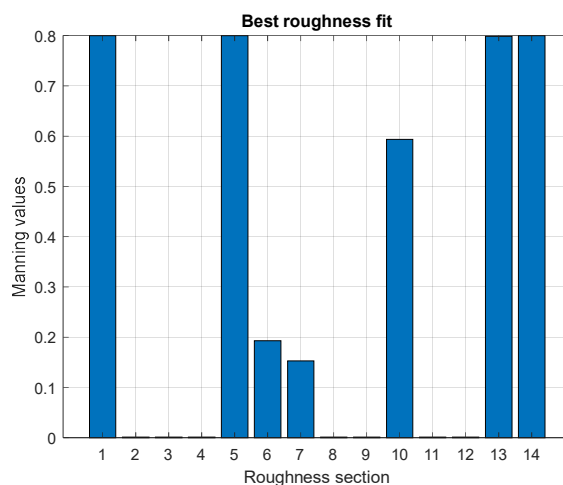
Figure 70: Roughness values after optimization
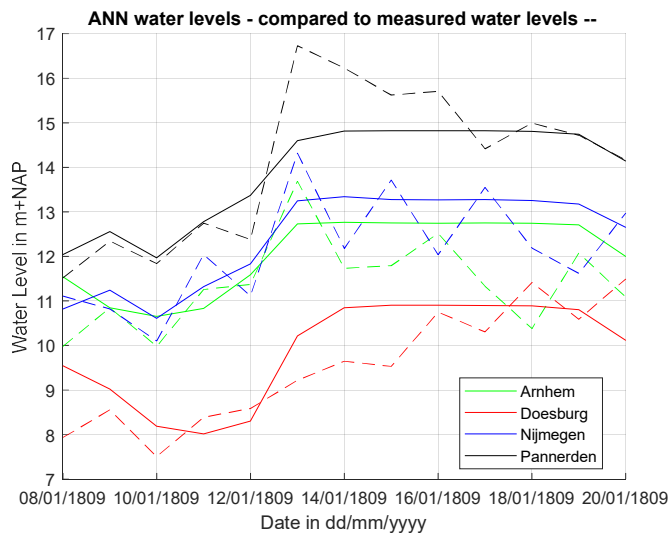
Figure 71: Water levels after optimization

9   Discussion

The first results show that the NARX is a good fit to the original flexible mesh model. This means that the NARX should represent the flexible mesh results in the same manner. Therefore, being within the roughness inputs and discharge waves the NARX should theoretically provide the correct results for all locations. This shows in having a fit of an R^2 value higher than 0.7 for all stations. This means that the overall course of the water level is correctly being displayed by the NARX.

The fit of the NARX does however vary between the different measuring stations, in which with roughness values Doesburg has the best fit and Nijmegen the worst fit. This can be explained by the amount of roughness sections that influence these measurement stations. The station at Doesburg, has relatively few other sections included to Arnhem and Pannerden. While the fit is being made over all stations, the other nodes, including feedback, will affect this accuracy. Having identical input parameters for other outputs will make these inputs have a stronger effect on the hidden layer. This results in these features being more presented in the output signal. For the case of this surrogated model, the locations Arnhem, Pannerden, and Nijmegen all have overlapping roughness sections that effect the water levels. Therefore, this roughness section is more present in the hidden layer as such the accuracies of fit of the values using this section will be better than the ones missing this section.

So, the fit is accurate however it remains an estimation and does not provide the same values. As errors progress through the calculation the eventual error is larger than the initial fitting error. This can cause issues with fitting the discharge wave to the measured water levels. One of the potential causes of the failing fits, is the progressions of these errors.

The fitting of the discharge wave to the measured values in 1809, has seen a lot of issues. These have several potential causes. First, the small error of the surrogated model fit leads to a larger error with values closer to the boundary of the sphere of fitting. Second, the amount of independent data points on which the error is determined is low. This is due to the close measurement stations together. To prevent this from being a problem, data from further measuring stations could be included or a longer period around the flood event could be considered. However, this will increase the time needed for the creation of the trainings data. Third, the roughness sections defined create a lot of unknown variables, more than independent data points.

The amount of data that is available in the measurements is limited. This is due to the historic nature of the case. In the past monitoring flows and water levels has been limited. This is increased a lot in the last 150 years. For the 1809 case, there are four measurement stations all located within 30km of each other, each of these stations has a daily water level measurement. However, due to the nature of river flows, the water level is persistence in both time and space. This means that the amount of independent measurements is lower than the total data available. In total there is an event of 13 days present with 4 stations, meaning that 52 data points are available. Considering the persistence of the system, this is most likely reduced, as the next day depends on the previous day and all station are close together. The amount of really independent data points remains unknown; however, it is probably less than the 16 unknown variables for the optimization problem. This means that the problem is badly defined, and as such can behave unexpected.

The optimization algorithm used has several boundary criteria. These criteria and the function values should be continuous in their value and first derivative to ensure the workings. In this problem, this direct requirement might not be met in all cases. As the dike breaches cause sudden changes in the behaviour of the input parameters. For the surrogated model this is implemented using a logical array which is discontinues. However, this is taken as a constant input value for the algorithm, so this

parameter does not break the criteria. Still it might lead to issues, that the first derivative of the output becomes discontinues when the dike breaches.

The measured values appear to be outside the sphere of fitting as such the surrogated model is no longer guaranteed valid. This could be one of the reasons that the optimization algorithm fails as the surrogated model does not represent the correct relation between the input and output parameters. So, the discharge wave might be incorrectly related to the water levels in this case.

10 Conclusion

The effectiveness of surrogated models for the reconstruction of historic discharges is extremely limited. The surrogated model used in this case a NARX, can represent the physical model. This is a physical model including dike breaches on a temporal resolution of 24 hours. The fit quality of the NARX to the physical model is good with the worst example being an R^2 of 0.75 and a maximum absolute deviation of 29cm. This deviation is a single outlier and overall, the results follow more closely. As such the NARX is a great tool to represent the physical model. Theoretically this should mean that the NARX can be used to change and test several inputs to meet a specific output.

However, it seemed that this changing the input parameters to meet a specific output is not working. Some reasons for this discrepancy are the low amount of independent data, the sphere of fitting created for the surrogated model, and the optimization algorithm used. One of these three problems is the cause of the optimization problem. More research is required to find the one that is causing the problems.

Currently the number of parameters that the NARX can use namely 19 is way more than the availability of independent data. The data must be independent as dependent data gives false relations and easier correlations by just the data structure itself. The number of data points available is 13 for each station, this is one per day. Of the 19 variables, 4 are fixed, which brings the number of unknown variables to 15, still more than the available direct data. The spatial correlation between the measurement stations is also highly present as they are within 30 km of each other. This means that the amount of available data is less than optimal. The optimization problem as such is poorly defined, which means that the number of functions is larger than the number of constrains/data.

The sphere of fitting is defined in only the theoretical roughness values and several discharge waves, this makes the range of potential roughness values limited. The roughness value is usually used to calibrate the hydrodynamic model. In the situation of historic floods this is impossible as to much data is missing. However, due to this property are there no clear limitations in roughness values except for being positive. This makes it difficult to create a training set including all potential roughness values.

The optimization algorithm used will try to find a global minimum, however no method can guarantee to reach a global minimum. This method uses evolutionary approach meaning that multiple start locations are chosen from which a steepest gradient search is undertaken. This approach has some limitations as the functions need to be continuous and differentiable. With the dike breaches, this second part might be an issue as the behaviour changes at a single point in time. This could make the derivative become discontinuous, leading to a failure in the algorithm.

Concluding a NARX is a powerful tool, in fitting the physical model. However, it has large limitations when it comes to output manipulation via the input. This shows in the algorithm trying to change the inputs to meet the measurements. This can also be caused by the algorithm itself however. As such more research is required to find the limitations of the NARX and the areas for which it can be used.

11 Recommendations for further research

The conclusions show that multiple factors are likely responsible for the bad results of the reconstruction. For this more research is advised in the areas of these three problems.

First, the ratio between known and unknown variables. This can be achieved by reducing the number of roughness sections. Next to this the length of the timeseries can be increased to get more independent data points. This could lead to a better representation of the problem, making the problem better defined, instead of having falsely defined relations or undefined relations. The falsely defined relations are relation using dependent data.

Second, the sphere of fitting can be enlarged by using a wider set of roughness values, this should allow for a wider range of values in which the surrogated model is operational. This can be achieved by creating a more diverse set of runs in DFlow-FM. The number of runs needed for this change, will dependent on the ability of the NARX to fit to this wider range of data, but it will most likely be equal or larger than 32, with the same 60/20/20 division used in this thesis.

Third, the optimization algorithm can be changed. This will probably cause the biggest difference in the workings of the system, while the NARX fits well to the original data. The importance for this system is to find an algorithm with less limitations. These are scares as most ill-defined problems are unsolvable. This is the other part in which an answer could be searched, to reduce the search space creating a better-defined problem. The use of non-gradient methods would remove the limitations on the derivatives but would increase computational power needed immensely. R

12 Acknowledgments

13 References

Al-khayyal, F. A., & Sherali, H. D. (2000). An Interior Point Algorithm for Large-scale Nonlinear Programming. *Society for Industrial and Applied Mathematics*, *10*(4), 1049–1057.

Alterra. (2004). Historisch Grondgebruik Nederland - WUR. Retrieved September 6, 2018, from https://www.wur.nl/nl/Onderzoek-Resultaten/Onderzoeksinstituten/Environmental-Research/Faciliteiten-Producten/Kaarten-en-GIS-bestanden/Historisch-Grondgebruik-Nederland.htm

Beeldbank (Gelders Archief) - Gelders Archief. (n.d.). Retrieved February 1, 2018, from https://www.geldersarchief.nl/bronnen/foto-s-en-films?mivast=37&miadt=37&mizig=284&miview=gal&milang=nl&micols=1&misort=last_mod%7Cdesc&mires=0&mizk_alle=doorbraken&mibj=1800&miej=1850

Behzad, M., Asghari, K., Eazi, M., & Palhang, M. (2009). Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Systems with Applications*, *36*(4), 7624–7629. https://doi.org/10.1016/j.eswa.2008.09.053

Bomers, A., Schielen, R. M. J., & Hulscher, S. J. M. H. (2018a). Hydraulic sensitivity analysis with a lower-fidelity surrogate model for the 1926 river Rhine flood event. *Manuscript Submitted for Publication*.

Bomers, A., Schielen, R. M. J., & Hulscher, S. J. M. H. (2018b). The influence of grid shape and grid size on hydraulic river modelling performance. *Manuscript Submitted for Publication*, 1–24.

Byrd, R. H., Gilbert, J. C., & Nocedal, J. (2000). A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming, Series B*, *89*(1), 149–185. https://doi.org/10.1007/PL00011391

Caviedes-Voullième, D., Morales-Hernández, M., López-Marijuan, I., & García-Navarro, P. (2014). Reconstruction of 2D river beds by appropriate interpolation of 1D cross-sectional information for flood simulation. *Environmental Modelling and Software*, *61*, 206–228. https://doi.org/10.1016/j.envsoft.2014.07.016

Cheng, B., & Titterington, D. M. (1994). Neural Networks: A Review from a Statistical Perspective. *Statistical Sciences*, *9*(1), 2–30. https://doi.org/10.1214/ss/1177010638

Chow, V. te. (1959). Open Channel Hydraulics. *Open Channel Hydraulics*.

Deltares. (2016a). D-Flow Flexible Mesh, Technical Reference Manual. Version: 1.1.0. Revision: 48665.

Deltares. (2016b). D-Flow Flexible Mesh, User Manual. Version 1.2.1.

Driessen, A. M. A. J. (1994). *Watersnood tussen Maas en Waal overstromingsrampen in het rivierengebied tussen 1780 en 1810*.

Duong, T. A., Bui, M. D., & Rutschmann, P. (2018). Data pre-processing combined with Artificial Neural Network to improve the performance of time series modelling, 659–660. https://doi.org/10.3850/978-981-11-2731-1

Dutch Ministry of Infrastructure and the Environment. (1926). Verslag van het voorgevallene tijdens het hooge opperwater op de Nederlandsche rivieren in den winter van 1925 op 1926. Retrieved from https://repository.tudelft.nl/islandora/object/uuid%3A5aee4476-3245-42da-a9aa-5807f2b77b6e

Dutch Ministry of Infrastructure and the Environment, & Ministry of Economic Affairs. (2014). Delta

Programme 2015 - Working on the delta - The decisions to keep the Netherlands safe and liveable, 180. Retrieved from https://english.deltacommissaris.nl/delta-programme/documents/publications/2014/09/16/delta-programme-2015

Ewijk, H. (1809a). Figuratieve kaart van de doorgebroken dijk onder Oosterholt in het ambt van Overbetuwe voorgevallen den 15 van Louwmaand 1809 voormiddag ten 9 uur. Amsterdam: Figuratieve kaart van de doorgebroken dijk onder Oosterholt in het ambt van Overbetuwe voorgevallen den 15 van Louwmaand 1809 voormiddag ten 9 uur. Retrieved from https://preserve2.archieven.nl/mi-37/fonc-gea/338e97cb8d5fe56af6bb513e4f67d3fe?format=thumb&miadt=37&miahd=779827702&mivast=37

Ewijk, H. (1809b). Figuratieve kaart van den doorgebroken dijk onder Loenen in het Ambt van Overbetuwe voorgevallen den 15 van Louwmaand 1809. Amsterdam: Joh. Allart + Jac. Rusch. Retrieved from https://preserve2.archieven.nl/mi-37/fonc-gea/4a9505b7b8a59a42c1986532b04c0c61?format=thumb&miadt=37&miahd=779827703&mivast=37

Ewijk, H. (1809c). Figurative kaart van den doorbraakte Loo in het ambt Lymers voorgevallen den 13 jan[uari] 1809. Amsterdam: Joh. Allart + Jac. Rusch. Retrieved from https://preserve2.archieven.nl/mi-37/fonc-gea/5e61419b1d0cc21aa5ceb25e38d21213?format=thumb&miadt=37&miahd=779827857&mivast=37

Hagan, M. T., Demuth, H. B., & Beale, M. H. (1995a). Neural Network Design. *Boston Massachusetts PWS*, *2*, 734. https://doi.org/10.1007/1-84628-303-5

Hagan, M. T., Demuth, H. B., & Beale, M. H. (1995b). Neural Network Design. *Boston Massachusetts PWS*, *2*, 734. https://doi.org/10.1007/1-84628-303-5

Hesselink, A. W., Stelling, G. S., Kwadijk, J. C. J., & Middelkoop, H. (2003). Inundation of a Dutch river polder, sensitivity analysis of a physically based inundation model using historic data. *Water Resources Research*, *39*(9), 1–17. https://doi.org/10.1029/2002WR001334

Kolkman, M. J., Kok, M., & van der Veen, A. (2005). Mental model mapping as a new tool to analyse the use of information in decision-making in integrated water management. *Physics and Chemistry of the Earth*, *30*(4–5 SPEC. ISS.), 317–332. https://doi.org/10.1016/j.pce.2005.01.002

Koziel, S., Ciaurri, D. E., & Leifsson, L. (2011). Surrogate-based methods. *Studies in Computational Intelligence*, *356*, 33–59. https://doi.org/10.1007/978-3-642-20859-1_3

Lintsen, W. (2009). Geschiedenis van de techniek in Nederland . De wording van een moderne samenleving 1800-1890 . Deel II Gezondheid en openbare hygiëne . Waterstaat en.

Matta, E., Duda, R., Scheer, C., Ma, Y., Zhang, Q., Hassan, A., … Schmid, A. (2018). Water level predictions using neural networks in critical gauges of the Rhine River , Germany. In *Proc. of the 5th IAHR Europe Congress — New Challenges in Hydraulic Research and Engineering* (pp. 657–658).

Nederlandsch Aardrijkskundig Genootschap. (1837). *Register der peilingen, behoordende tot de kaart der rivieren de boven-rijn, de Waal, de Merwe, de oude- en een gedeelte der nieuwe-Maas van Lobith tot Brielle.*

Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Kevin Tucker, P. (2005). *Surrogate-based analysis and optimization. Progress in Aerospace Sciences* (Vol. 41). https://doi.org/10.1016/J.PAEROSCI.2005.02.001

Raman, H., & Chandramouli, V. (1996). Deriving a General Operating Policy for Reservoirs Using Neural Network. *Journal of Water Resources Planning and Management*, *122*(5), 342–347. https://doi.org/10.1061/(ASCE)0733-9496(1996)122:5(342)

Razavi, S., Tolson, B. A., & Burn, D. H. (2012a). Numerical assessment of metamodelling strategies in computationally intensive optimization. *Environmental Modelling and Software*, *34*, 67–86. https://doi.org/10.1016/j.envsoft.2011.09.010

Razavi, S., Tolson, B. A., & Burn, D. H. (2012b). Review of surrogate modeling in water resources. *Water Resources Research*, *48*(7). https://doi.org/10.1029/2011WR011527

Rijkswaterstaat. (n.d.). Parameters - Bulkdownload - Rijkswaterstaat Waterinfo. Retrieved September 7, 2018, from http://waterinfo.rws.nl/#!/bulkdownload/locatie-lijst/

Rijkswaterstaat. (1994). *Tienjarig overzicht 1981-1990*.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., … Tarantola, S. (2008). *Global sensitivity analysis: the primer*.

Simpson, T. W., Peplinski, J. D., Koch, P. N., & Allen, J. K. (2001). Metamodels for computer-based engineering design: Survey and recommendations. *Engineering with Computers*, *17*(2), 129–150. https://doi.org/10.1007/PL00007198

Toonen, W. H. J. (2013). *A Holocene flood record of the Lower Rhine*. Retrieved from http://dspace.library.uu.nl/handle/1874/282871

Appendix A: All accuracy graphs discharge only

Run	Set 1	Set 2
1		
2		
3		

Run	Set 1	Set 2
4		
5		
6		

Run	Set 1	Set 2

7

8

9

Run	Set 1	Set 2
10		
11		
12		

Run	Set 1	Set 2
13		
14		
15		
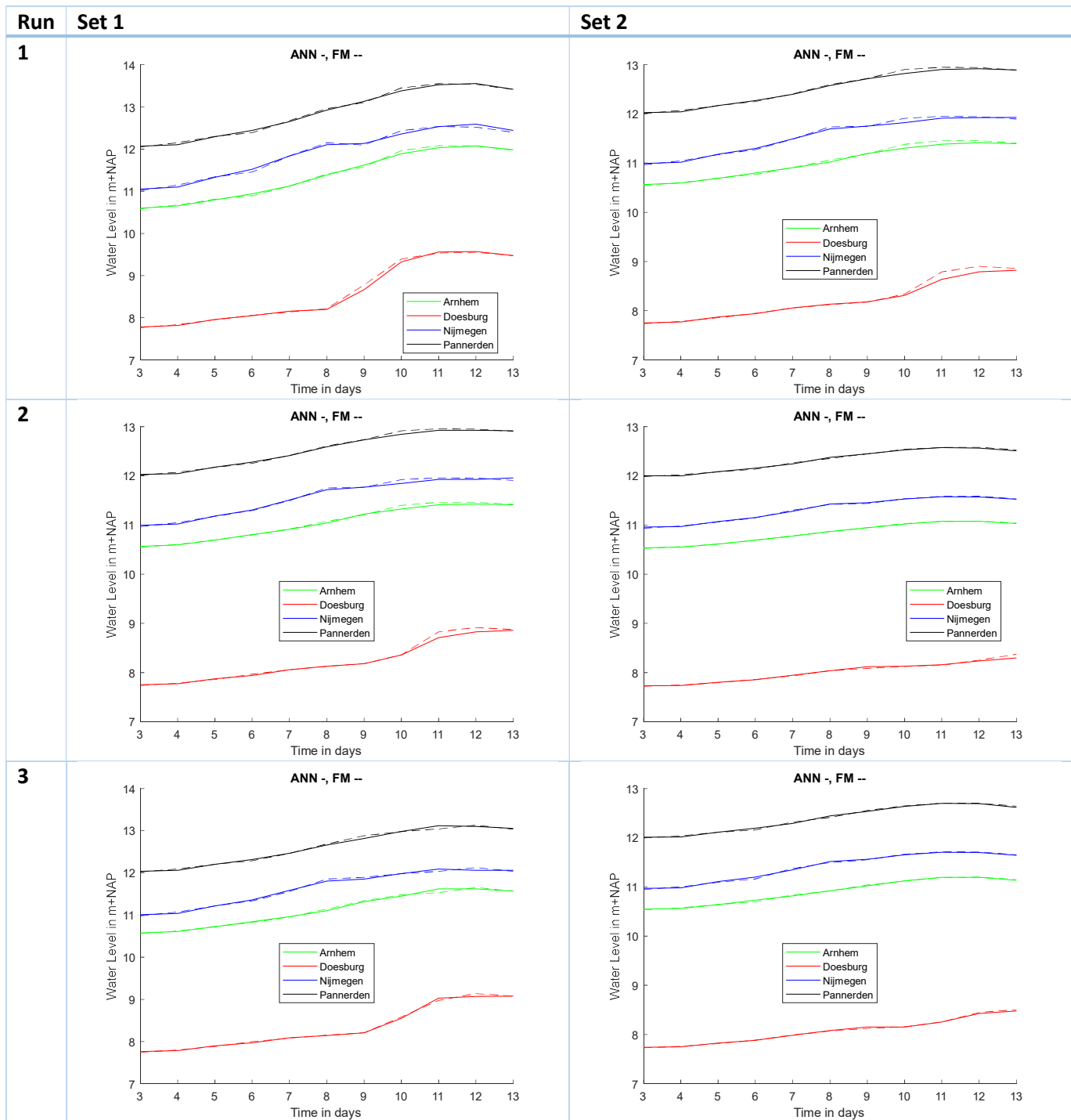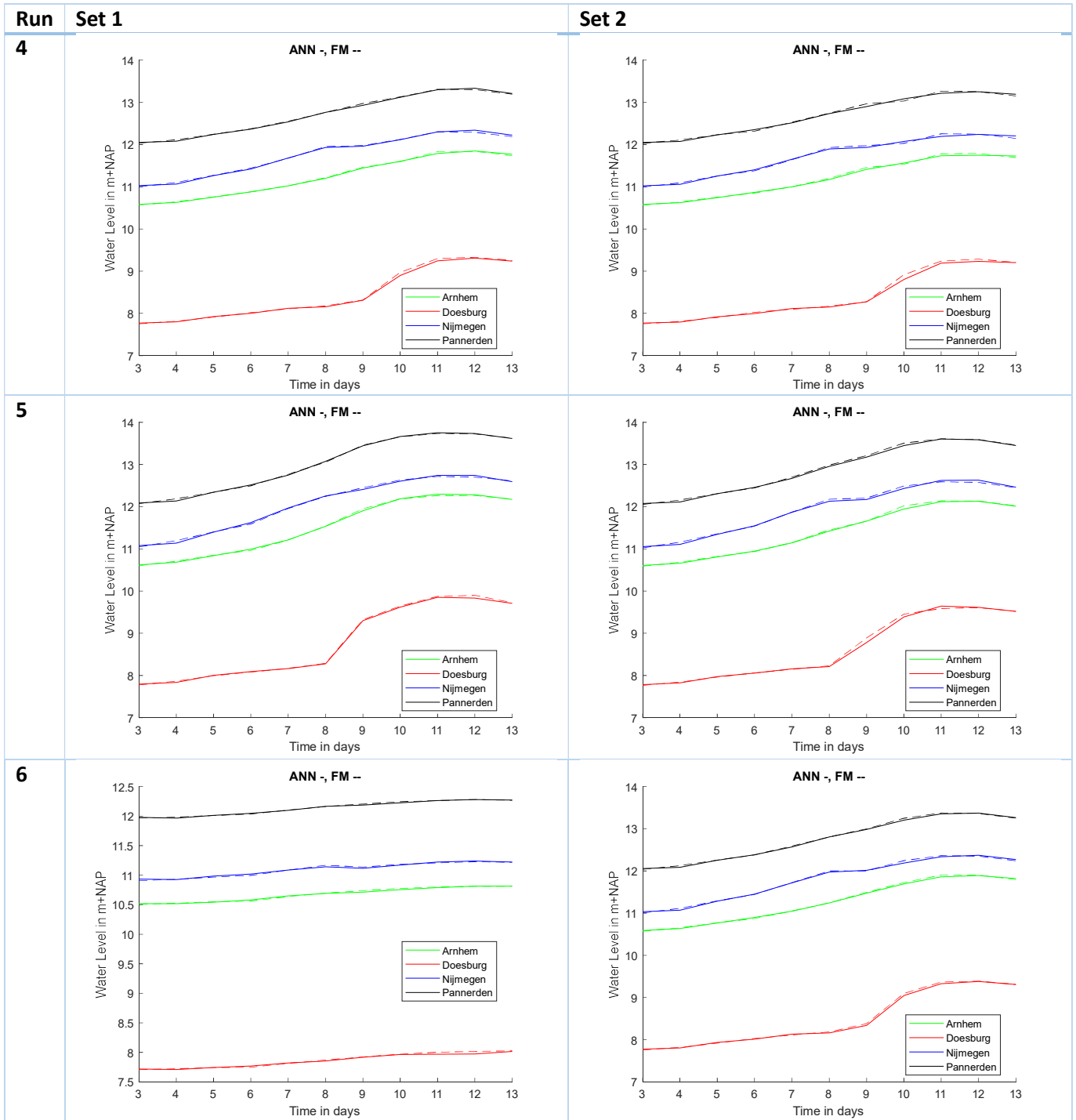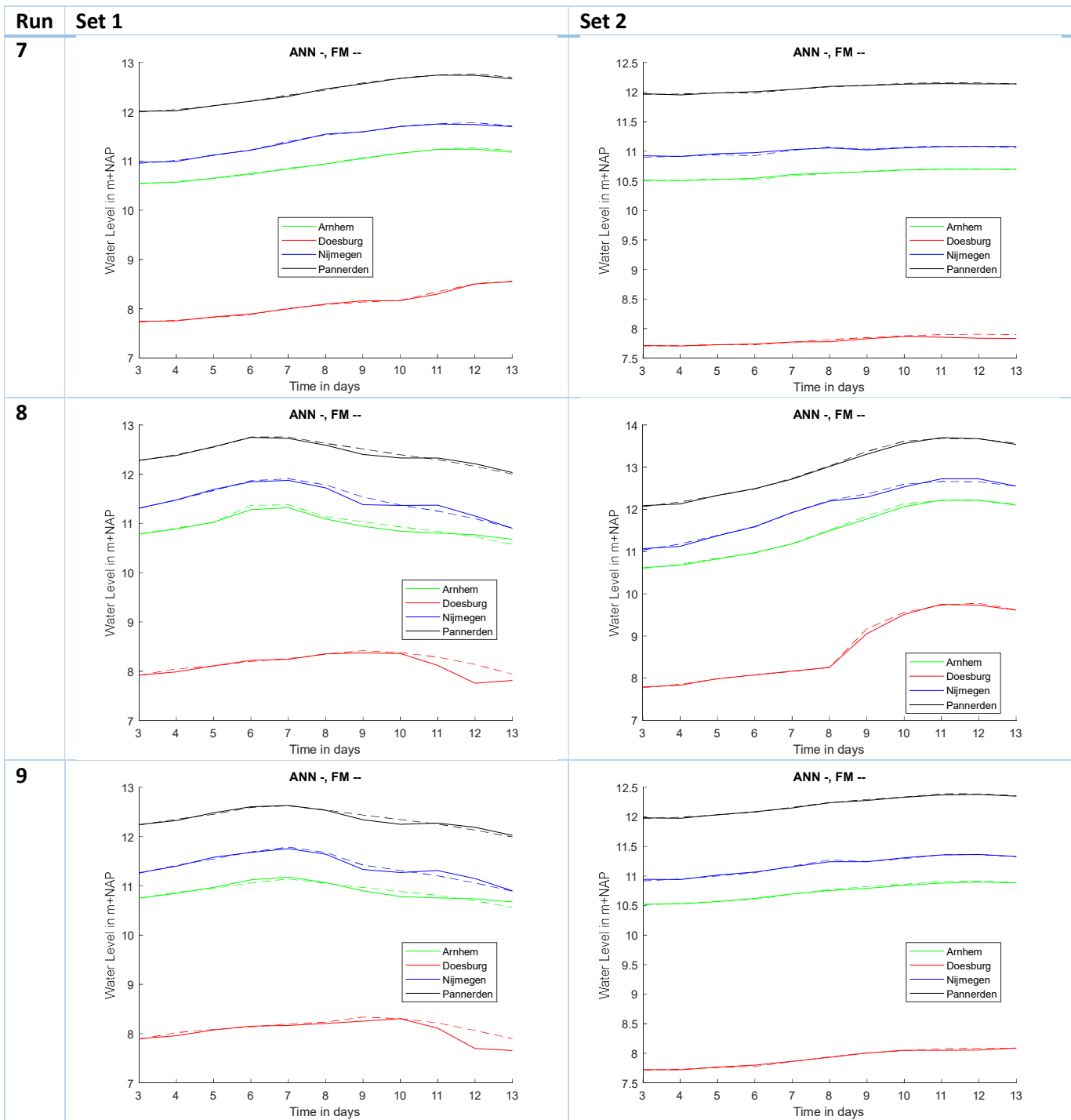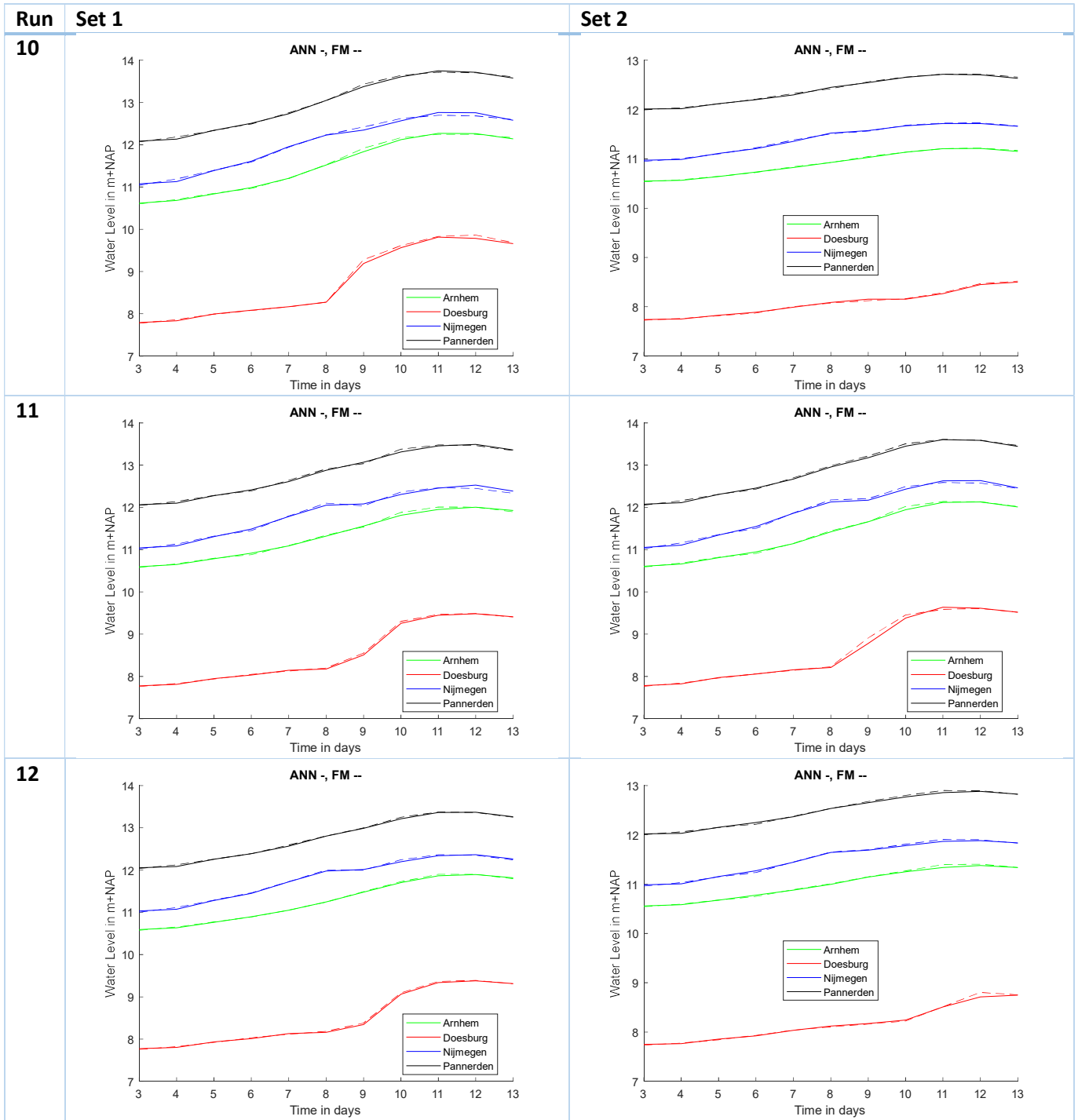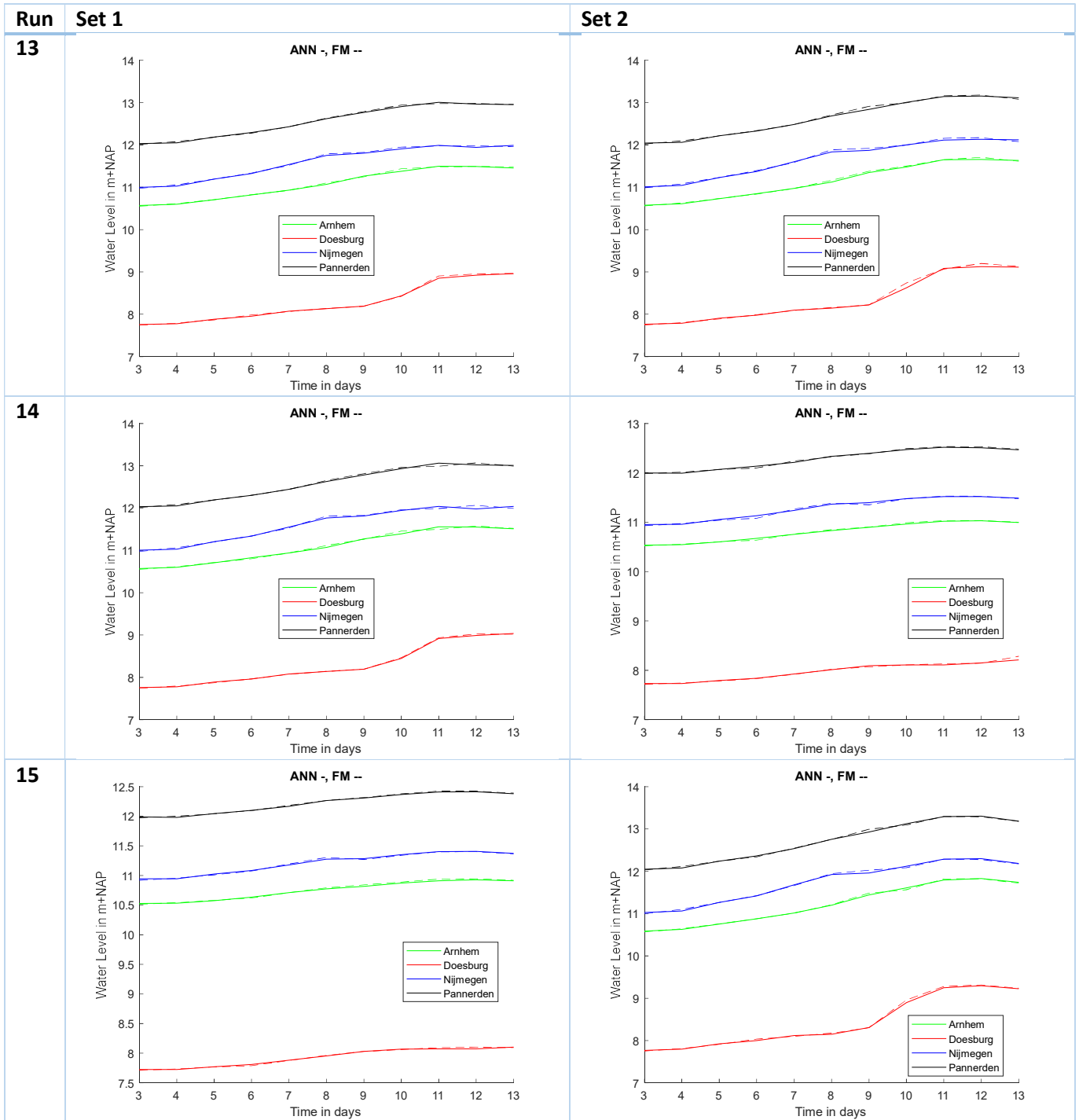
Run	Set 1	Set 2
16		
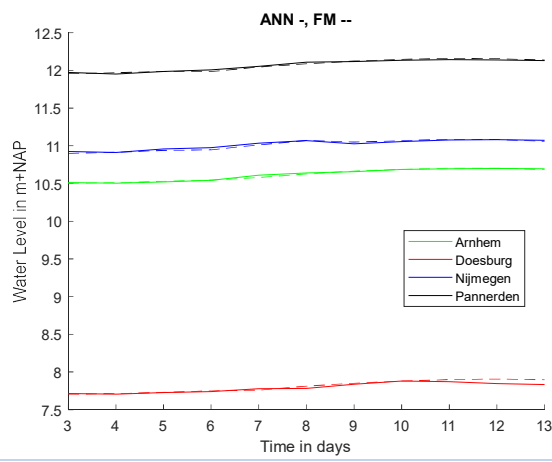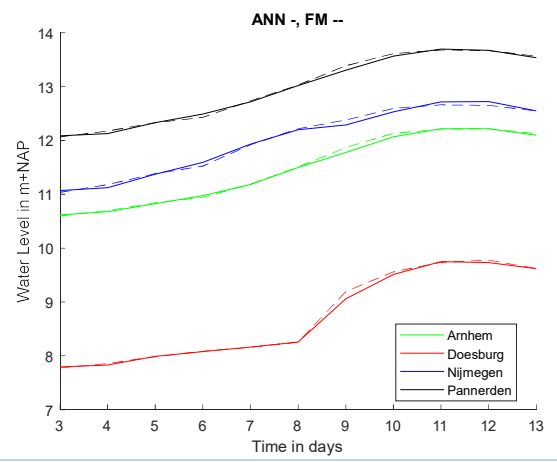
Appendix B: All accuracy graphs discharge and roughness

Run	Set 1	Set 2
1		
2		
3		

16

ANN -, FM --

ANN -, FM --